



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 163902**

**TO: Delia Ramirez**  
**Location: 2d74 / 2c70**  
**Art Unit: 1652**  
**Monday, August 29, 2005**

**Case Serial Number: 09371347**

**From: Noble Jarrell**  
**Location: Biotech-Chem Library**  
**Rem 1B71**  
**Phone: 272-2556**

**Noble.jarrell@uspto.gov**

### **Search Notes**

**This Page Blank (uspto)**





|||||  
Db 320 CTGCGGTTATTTCTTGTCTCACTGCGGTATGGTTACTGGGTCTCGGTGATTGAGAA 379  
Qy 301 TACACCTACTTTTGCAATGGGGGAGATTAATTGATTAAGACTTCAAGAGCTTGGAGCC 360  
Db 380 TACACCTACTTTTGCAATGGGGGAGATTAATTGATTAAGACTTCAAGAGCTTGGAGCC 439  
Qy 361 CGGATTTCTATGACACTGACATGACATGACTGTGTAGAGTTTGAACCTGTGTGAG 420  
Db 440 CGGATTTCTATGACACTGACATGACATGACTGTGTAGAGTTTGAACCTGTGTGAG 499  
Qy 421 CCGTGATTTGCTGACCTGTGGCCAGCCCTCAAGAAAGATTTTATGCTCAAGAGAGCAA 480  
Db 500 CCGTGATTTGCTGACCTGTGGCCAGCCCTCAAGAAAGATTTTATGCTCAAGAGAGCAA 559  
Qy 481 GAGGAGTAAAGTGGCGACTCCGGGTGGCAATCTGATCTGCTTGAAGAGAGAGCTTGTG 540  
Db 560 GAGGAGTAAAGTGGCGACTCCGGGTGGCAATCTGATCTGCTTGAAGAGAGAGCTTGTG 619  
Qy 541 AAGTCAGAGCTGTACATCAATTTGATCTCAAGTCTGAGCTTCTGAGATTGATGATTCAGGA 600  
Db 620 AAGTCAGAGCTGTACATCAATTTGATCTCAAGTCTGAGCTTCTGAGATTGATGATTCAGGA 679  
Qy 601 AGAAAGATTTCTGAGGTTTGAAGCAAAATGCAAGTGAACCAATCCATGTTGTA 660  
Db 680 AGAAAGATTTCTGAGGTTTGAAGCAAAATGCAAGTGAACCAATCCATGTTGTA 739  
Qy 661 ATTGAAGATTTGAGCTCTCACTTACCCGTTCCGTTACCCCACTCTCAAGAGCTCTCTG 720  
Db 740 ATTGAAGATTTGAGCTCTCACTTACCCGTTCCGTTACCCCACTCTCAAGAGCTCTCTG 799  
Qy 721 AATATTCCTGGTTTACCCCGAATATTTACAGGTATCTGAGAGTCTCTTGGCCAG 780  
Db 800 AATATTCCTGGTTTACCCCGAATATTTACAGGTATCTGAGAGTCTCTTGGCCAG 859  
Qy 781 GAGGAAAGCCAAAGTATCTGTGACTTCAAGCAATCTGATTTTCAAGTCCCAATTTCAAAG 840  
Db 860 GAGGAAAGCCAAAGTATCTGTGACTTCAAGCAATCTGATTTTCAAGTCCCAATTTCAAAG 919  
Qy 841 GCAATTCATCTTACGAAATGATGCAATAAACAATCTGCTGTGATGATTTGACATT 900  
Db 920 GCAATTCATCTTACGAAATGATGCAATAAACAATCTGCTGTGATGATTTGACATT 979  
Qy 901 TCAAAATACAGCTTTCTATCACTGAGAGTCTTCAAGCGTATCTGCGCTTAACAGT 960  
Db 980 TCAAAATACAGCTTTCTATCACTGAGAGTCTTCAAGCGTATCTGCGCTTAACAGT 1039  
Qy 961 GATTTGAGATGCAAGCTTATCTCAAGAGTGAAGTGAAGTAAAGAGAGCACTGCG 1020  
Db 1040 GATTTGAGATGCAAGCTTATCTCAAGAGTGAAGTGAAGTAAAGAGAGCACTGCG 1099  
Qy 1021 GTCCTTTGAAATTAAGGAGAGAGAGAGAGAGTACTTACCCAGACATATA 1080  
Db 1100 GTCCTTTGAAATTAAGGAGAGAGAGAGAGAGTACTTACCCAGACATATA 1159  
Qy 1081 CTTGCGGAGATTTCTCTCAAGTCAATTTTACCTGTGTGTTGAATTCGAGCAATTCCT 1140  
Db 1160 CTTGCGGAGATTTCTCTCAAGTCAATTTTACCTGTGTGTTGAATTCGAGCAATTCCT 1219  
Qy 1141 AAAAAGCATTTTGGGAGCCCTTGTGACATATACAGTGAAGTGTGAAAAAGCGAGG 1200  
Db 1220 AAAAAGCATTTTGGGAGCCCTTGTGACATATACAGTGAAGTGTGAAAAAGCGAGG 1279  
Qy 1201 CTACAGAGGCTGTGACATTAACAGAGGAGCGAGATTAAGCGCTTGTGACAGATGCC 1260  
Db 1280 CTACAGAGGCTGTGACATTAACAGAGGAGCGAGATTAAGCGCTTGTGACAGATGCC 1339  
Qy 1261 TGTGCTGTGTTGATCTCTCTCTGCTTCTTCTTGGCAGCAACAATCAAGTCTC 1320  
Db 1340 TGTGCTGTGTTGATCTCTCTCTGCTTCTTCTTGGCAGCAACAATCAAGTCTC 1399  
Qy 1321 CTGCTCAACATCTTCTTAACTTCAACCGAGACATATTTGTGTGCAAGTCAAGTTTA 1380  
|||||

Db 1400 CTGCTCAACATCTTCTTAACTTCAACCGAGACATATTTGTGTGCAAGTCAAGTTTA 1459  
Qy 1381 TTTCACCCAGAAAGCTCCATTTTGTCTTCAACATTTGAGAAATTTCTGCTACTGACACA 1440  
Db 1460 TTTCACCCAGAAAGCTCCATTTTGTCTTCAACATTTGAGAAATTTCTGCTACTGACACA 1519  
Qy 1441 ACAGAGTTCTGCGAAGGAGATATGATACAGCTGCGCTTGTGTGTTGCTTCAAGTT 1500  
Db 1520 ACAGAGTTCTGCGAAGGAGATATGATACAGCTGCGCTTGTGTGTTGCTTCAAGTT 1579  
Qy 1501 CTTGACCCAAACATATATGATATCCCATGAAAGAGAGGAGAAAGCCCTGCTCTTAAGATA 1560  
Db 1580 CTTGACCCAAACATATATGATATCCCATGAAAGAGAGGAGAAAGCCCTGCTCTTAAGATA 1639  
Qy 1561 TCCATCTCTCTGCAACAACAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
Db 1640 TCCATCTCTCTGCAACAACAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1699  
Qy 1621 ATATGCTGCTGCAAGAAACCGGCAATAGCCCGTTTATTTGGTTCTTACACATAGAGAG 1680  
Db 1700 ATATGCTGCTGCAAGAAACCGGCAATAGCCCGTTTATTTGGTTCTTACACATAGAGAG 1759  
Qy 1681 AAATCCCAAGAAACAACCCAGATGAAATTTTGGAGCAATGTTGTTTGGCTGC 1740  
Db 1760 AAATCCCAAGAAACAACCCAGATGAAATTTTGGAGCAATGTTGTTTGGCTGC 1819  
Qy 1741 AGGCAATAGGATAGGATTAATCTATTCAGAAAGAGCTCAGACATTTCTTAAAGATGGG 1800  
Db 1820 AGGCAATAGGATAGGATTAATCTATTCAGAAAGAGCTCAGACATTTCTTAAAGATGGG 1879  
Qy 1801 ATCTTAACTCATTAAGGTTTCTTCTCAAGAGATGCTCTGTGTGGAGAGAGAGAGCC 1860  
Db 1880 ATCTTAACTCATTAAGGTTTCTTCTCAAGAGATGCTCTGTGTGGAGAGAGAGAGCC 1939  
Qy 1861 CCAAGAAAGTATGATACAGCAACATCCAGCTTCAATGAGGAGAGGAGAGAGATCTC 1920  
Db 1940 CCAAGAAAGTATGATACAGCAACATCCAGCTTCAATGAGGAGAGGAGAGAGATCTC 1999  
Qy 1921 CTCACAGAGAAACGGCATATTTATGTGTGAGATGCAAGAAATATGSCCAAGATGTA 1980  
Db 2000 CTCACAGAGAAACGGCATATTTATGTGTGAGATGCAAGAAATATGSCCAAGATGTA 2059  
Qy 1981 CATGATGCCCTTGTGCAATTAATAGCAAGAGAGTGTGAGTGAATAAATCTAGAAAGCATG 2040  
Db 2060 CATGATGCCCTTGTGCAATTAATAGCAAGAGAGTGTGAGTGAATAAATCTAGAAAGCATG 2119  
Qy 2041 AAAACCTGGGCACTTTAAAGAAAGAAAGCGTACCTCAGGATATTTGTCTATAA 2097  
Db 2120 AAAACCTGGGCACTTTAAAGAAAGAAAGCGTACCTCAGGATATTTGTCTATAA 2176  
|||||

RESULT 2  
AX050463 3259 bp DNA linear PAT 12-JAN-2001  
LOCUS AX050463  
DEFINITION Sequence 23 from Patent WO0071754.  
ACCESSION AX050463  
VERSION AX050463.1 GI:12226668  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 Johnson, W.G. and Stenroos, B.S.  
METHODS for diagnosing, preventing, and treating developmental  
disorders due to a combination of genetic and environmental factors  
Patent: WO 0071754-A 23 NOV-2000;  
JOURNAL University of Medicine and Dentistry of New Jersey (US)  
location/Qualifiers  
FEATURES  
1..3259  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

**ORIGIN**

100.0%; Score 2097; DB 6; Length 3259;

Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2097; Conservative 0; Indels 0; Gaps 0;

QY	1	ATGAGGAGGTTCTGTTACTATATATGCTACACAGGAGGGGACAGGAAAGGCCATGGCAGAA	60
Db	80	ATGAGGAGGTTCTGTTACTATATATGCTACACAGGAGGGGACAGGAAAGGCCATGGCAGAA	139
QY	61	GAAATGTGTGAGCAAGCTGTGTACATGATATTTCTGCAGATCTTCACTGTATTAATGCAA	120
Db	140	GAAATGTGTGAGCAAGCTGTGTACATGATATTTCTGCAGATCTTCACTGTATTAATGCAA	199
QY	121	TCCGATTAAGTATGACTTAATAAACCGAAACAGCTCCTCTTGTGTGTGTGTTCTTACACG	180
Db	200	TCCGATTAAGTATGACTTAATAAACCGAAACAGCTCCTCTTGTGTGTGTGTTCTTACACG	259
QY	181	GGCACCCGAGAGCCCAACCCGACACAGCCCGGCAAGTTGTTAAGGAAATACAGAACCAACA	240
Db	260	GGCACCCGAGAGCCCAACCCGACACAGCCCGCAAGTTGTTAAGGAAATACAGAACCAACA	319
QY	241	CTGC CGGTGTAATTTCTTGTGCTCACCTGCGATATGAGTTACTGGGTCCTGCGATATTCGAA	300
Db	320	CTGC CGGTGTAATTTCTTGTGCTCACCTGCGATATGAGTTACTGGGTCCTGCGATATTCGAA	379
QY	301	TACACCTTACTTTTGCAATGCGGGGAGATTAATTGATTAACGACTTCAAGACCTTGAAGCC	360
Db	380	TACACCTTACTTTTGCAATGCGGGGAGATTAATTGATTAACGACTTGAAGACCTTGAAGCC	439
QY	361	CGGCAATTTCTATGACATCTGACATGACAGATGACCTGTGATGGTTTAAAGACTTGTGGTTGAG	420
Db	440	CGGCAATTTCTATGACATCTGACATGACAGATGACCTGTGATGGTTTAAAGACTTGTGGTTGAG	499
QY	421	CCGTGATTTGCTGGAATCTGTGCGCAGCCCTCGAAGAACATTTTAAAGTCAAGCAGAGACA	480
Db	500	CCGTGATTTGCTGGAATCTGTGCGCAGCCCTCGAAGAACATTTTAAAGTCAAGCAGAGACA	559
QY	481	GAGAGATTAATGAGGCGCATCCCGGTGGCATCACTGATATCTTGAAGACAGACCTTGTG	540
Db	560	GAGAGATTAATGAGGCGCATCCCGGTGGCATCACTGATATCTTGAAGACAGACCTTGTG	619
QY	541	AAGTCAGAGCTGCTACACATATGAATCTCAAGTCGAGCTTCTGAAGATTGATGATTCAGGA	600
Db	620	AAGTCAGAGCTGCTACACATATGAATCTCAAGTCGAGCTTCTGAAGATTGATGATTCAGGA	679
QY	601	AGAAAGATTTCTGAGTTTGAAGCAAAATCAGTGAACAGCAACCAATCCAAATGTTGTA	660
Db	680	AGAAAGATTTCTGAGTTTGAAGCAAAATCAGTGAACAGCAACCAATCCAAATGTTGTA	739
QY	661	ATTGAAGACTTGAAGTCTCACTTACCCGTTGCGTATACCCCACTCTCAAGCCTCTCTG	720
Db	740	ATTGAAGACTTGAAGTCTCACTTACCCGTTGCGTATACCCCACTCTCAAGCCTCTCTG	799
QY	721	AATATTCCTGAGTTTACCCCGAAGATATTTACAGGTACATCTGACAGAGATCTCTTGGCCAG	780
Db	800	AATATTCCTGAGTTTACCCCGAAGATATTTACAGGTACATCTGACAGAGATCTCTTGGCCAG	859
QY	781	GAGGAAAGCCAGATATCTGTGACTTCAAGCAGATCCAGTTTTTCAAGTCCCAATTTCAAAG	840
Db	860	GAGGAAAGCCAGATATCTGTGACTTCAAGCAGATCCAGTTTTTCAAGTCCCAATTTCAAAG	919
QY	841	GCAGTTCAACTTACAGATGATGCCATTAATAAACACTCTGCTGTGTGATTTGGACATT	900
Db	920	GCAGTTCAACTTACAGATGATGCCATTAATAAACACTCTGCTGTGTGATTTGGACATT	979
QY	901	TCAATAACAGACTTTTCCATACGCTTGAGATGCTTACAGCGTATCTGCCTTAAACAGT	960
Db	980	TCAATAACAGACTTTTCCATACGCTTGAGATGCTTACAGCGTATCTGCCTTAAACAGT	1039
QY	961	GATTCGAGGATACAAAGCCTTACTCCAAAGCTGACGCTTGAAGATTAAGAAGAGACATGCG	1020
Db	1040	GATTCGAGGATACAAAGCCTTACTCCAAAGCTGACGCTTGAAGATTAAGAAGAGACATGCG	1099

QY	1021	GTCCCTTTGAAAAAATAAAGGCAACAAAGAAAGAGACTACCTTAACCCAGCAATTA	1080
Db	1100	GTCCCTTTGAAAAAATAAAGGCAACAAAGAAAGAGACTACCTTAACCCAGCAATTA	1159
QY	1081	CCTCGCGGAGTGTCTCTCCAGTTCAATTTTAACTGTGTCTTGAAATCCAGCAATTCCT	1140
Db	1160	CCTCGCGGAGTGTCTCTCCAGTTCAATTTTAACTGTGTCTTGAAATCCAGCAATTCCT	1219
QY	1141	AAAAAGCATTTTTCGAGGCCCTTGTSAGACTATACAGTGCACAGTGCATGAAGGCGAGG	1200
Db	1220	AAAAAGCATTTTTCGAGGCCCTTGTSAGACTATACAGTGCACAGTGCATGAAGGCGAGG	1279
QY	1201	CTACAGGAGCTGTGCAGTAAACAAGGGGACGCCATATATAGCCGCTTTGTACAGAGATGCC	1260
Db	1280	CTACAGGAGCTGTGCAGTAAACAAGGGGACGCCATATATAGCCGCTTTGTACAGAGATGCC	1339
QY	1261	TGTGCTGTGTGTGATCTCTCCCTGCTTCCCTCTTTCGACGACCACTCAGTCTC	1320
Db	1340	TGTGCTGTGTGTGATCTCTCCCTGCTTCCCTCTTTCGACGACCACTCAGTCTC	1399
QY	1321	CTGCTCCGAACTTTTCTTAACTTTCACCCAGACATATCTGTGTGCAAGCTCAAGTTTA	1380
Db	1400	CTGCTCCGAACTTTTCTTAACTTTCACCCAGACATATCTGTGTGCAAGCTCAAGTTTA	1459
QY	1381	TTTCAACCCAGGAAAGCTCCATTTTGTCTTCAACAATTGTGAAATTTCTGTACTACGCACA	1440
Db	1460	TTTCAACCCAGGAAAGCTCCATTTTGTCTTCAACAATTGTGAAATTTCTGTACTACGCACA	1519
QY	1441	ACAGAGGTTCTCGGAAAGGAGATATGTAACAGCTGCGCTGTGTGTGTTGCTTCAGTT	1500
Db	1520	ACAGAGGTTCTCGGAAAGGAGATATGTAACAGCTGCGCTGTGTGTGTTGCTTCAGTT	1579
QY	1501	CTTGAGCCAAACATACATGATGATCCCATGAGGAACAGCGGAGAAAGCCCTGAGCTCCATAGTA	1560
Db	1580	CTTGAGCCAAACATACATGATGATCCCATGAGGAACAGCGGAGAAAGCCCTGAGCTCCATAGTA	1639
QY	1561	TCCATCTCTCTCGAACAACAATTCCTTCCACTTACCAATGACCCCTCAATCCCAATC	1620
Db	1640	TCCATCTCTCTCGAACAACAATTCCTTCCACTTACCAATGACCCCTCAATCCCAATC	1699
QY	1621	ATAATGTGGGTCCAGGAACCGGATAGCCCGCTTATTTGGTTCCTACAAATATAGAGAG	1680
Db	1700	ATAATGTGGGTCCAGGAACCGGATAGCCCGCTTATTTGGTTCCTACAAATATAGAGAG	1759
QY	1681	AAACTCCAAAGAACAAACCAGATGGAATTTTSGAACAAATGTGTGTTTTTGGGTGCG	1740
Db	1760	AAACTCCAAAGAACAAACCAGATGGAATTTTSGAACAAATGTGTGTTTTTGGGTGCG	1819
QY	1741	AGGCATTAAGATAGGAGTATCTATTCAGAAAGAGCTCAGACATTTTCTTTAGCATGCG	1800
Db	1820	AGGCATTAAGATAGGAGTATCTATTCAGAAAGAGCTCAGACATTTTCTTTAGCATGCG	1879
QY	1801	ATCTTAATCTCATTAAGGTTTCTCTCAAGANAATCTCTGTGTGGGAGGAGAGAGCC	1860
Db	1880	ATCTTAATCTCATTAAGGTTTCTCTCAAGANAATCTCTGTGTGGGAGGAGAGAGCC	1939
QY	1861	CCAGCAAAAGTATGTACAAACAACATCAGCTTCATGAGCCAGCAGTGTCGAGATCTC	1920
Db	1940	CCAGCAAAAGTATGTACAAACAACATCAGCTTCATGAGCCAGCAGTGTCGAGATCTC	1999
QY	1921	CTCCAGAGAAAGGCGCATATTTTATGTGTGAGATGCAAAAGATATGTGCCAAGATGTA	1980
Db	2000	CTCCAGAGAAAGGCGCATATTTTATGTGTGAGATGCAAAAGATATGTGCCAAGATGTA	2059
QY	1981	CATGATGCCCTTGTGCAAAATATATAGAAAGAGGTTTGAAGTGAATACTTGAAGCATG	2040
Db	2060	CATGATGCCCTTGTGCAAAATATATAGAAAGAGGTTTGAAGTGAATACTTGAAGCATG	2119
QY	2041	AAAACTGTGCGCACTTTAAAGAGAAAAAGCTTACTCTCAGATATTTGTGCTATA	2097
Db	2120	AAAACTGTGCGCACTTTAAAGAGAAAAAGCTTACTCTCAGATATTTGTGCTATA	2176

RESULT 3	AF025794	3259 bp	mRNA	linear	PRI 26-MAR-1998
LOCUS	AF025794				
DEFINITION	AF025794				
ACCESSION	AF025794				
VERSION	AF025794.1				
KEYWORDS	GI:2981302				
SOURCE					
ORGANISM	Homo sapiens (human)				
REFERENCE	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 3259) Leclerc, D., Wilson, A., Dunas, R., Gafuick, C., Song, D., Watkins, D., Heng, H.H.Q., Kommens, J.M., Scherer, S.W., Rosenblatt, D.S. and Gravel, R.A. Cloning and mapping of a cDNA for methionine synthase reductase, a flavoprotein defective in patients with homocystinuria Proc. Natl. Acad. Sci. U.S.A. 95 (6), 3059-3064 (1998)				
TITLE					
JOURNAL					
MEDLINE	98169496				
PUBMED	9501215				
REFERENCE	2 (bases 1 to 3259) Leclerc, D. Direct Submission Submitted (19-SEP-1997) Human Genetics, McGill University - Montreal Children's Hospital Research Institute, 4060 Ste-Catherine West, Montreal, Que H3Z 2Z3, Canada 3 (bases 1 to 3259) Leclerc, D. Direct Submission Submitted (12-NOV-1997) Human Genetics, McGill University - Montreal Children's Hospital Research Institute, 4060 Ste-Catherine West, Montreal, Que H3Z 2Z3, Canada Sequence update by submitter Location/Qualifiers 1..3259 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /chromosome="5" /map="5p15.2-p15.3" 1..3259 /gene="MTRR" 80..2176 /gene="MTRR" /note="B12 vitamin; cblE complementation group" /product="methionine synthase reductase" /codon_start=1 /protein_id="AAC39667.1" /db_xref="GI:2981303" /translation="MRRPLILLYATGQGGAKAIAEMCEQAVYHGFSPADIHGISDDKY DKETRIAPLVVAVSTGTGGDDPPTARFVEIKQNTLPVDFFAHRRGULGDSYTT YFCNGKATIDKRLQELGARHYDYTGHADDCVGLVVEPMTAGLWPLRRHFRSSRQ EIRIGALPVASPAIRLTDLVKSELHIESQVELRPDPSGRKDSBVAKONAVNSQSN VVIDFESSLTSRVPVLSQASINLIGLPEYVLYQLQESLQESQSVISADVPQV PISAVQLTINDAIKTLTILVLDISNDPFSYQPDGASVICPNSDSQSLQRLQTE DKRHCYLTKKADTKKKKGAIRPHIAGCSLQIPFTWCLEIRAIIPKKAIRALVDYTT SDSABKRRLQELCSKQGAADYSRRVRDACAQLDLALFSCQPPSLILHLEKLPQ RPSYCASSLSFHPKELHVENIVEFLSTATTEVRKGVCTGMLLVAASVILQPIHNS HEDGKALAPKISISPTRTNSFHLDPDSIPINVGPGTGLAPETGLOHEKLEQBOH PDGNFGAMMLFFGCRHKRDYLFRELEHPLKGIILTKVSPRDAPVEEAPAKKX VDNIHLHGQGVARIILQENHITVCGDAKMAVDVHALVQIISKEVGEKLEAMKMT LATLKERKRYQIDWMS"				
REMARK					
FEATURES					
SOURCE					
gene					
CDS					
ORIGIN					
Query Match	100.0%;	Score 2097;	DB 9;	Length 3259;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches 2097;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
DB	1	ATGAGGAGTTCTTCTATATGCTACACAGCAGGAGCGGCAAGAGCCATGCCAGAA	60		
	80	ATGAGGAGTTCTTCTATATGCTACACAGCAGGAGCGGCAAGAGCCATGCCAGAA	139		

QY	61	GAATGTGTGAGCAAGCTGTGTGATCAATGTGATTTTCTGCAATCTTCACTATATTAGAA	120
Db	140	GAATGTGTGAGCAAGCTGTGTGATCAATGTGATTTTCTGCAATCTTCACTATATTAGAA	199
QY	121	TCCGATTAAGTAGACCTAAAAACCGAAACAGCTCTCTTGTGTGTTTCTTACACG	180
Db	200	TCCGATTAAGTAGACCTAAAAACCGAAACAGCTCTCTTGTGTGTTTCTTACACG	259
QY	181	GGCACCGGAGAACCCACCCGACACAGCCCGCAGTTTGTTAAGGAAATACAGAACCAACA	240
Db	260	GGCACCGGAGAACCCACCCGACACAGCCCGCAGTTTGTTAAGGAAATACAGAACCAACA	319
QY	241	CTGCCGGTTGATTTCTTTTGTCTCACCTGCGGTATGGTTACTGGGTCCTCGTGATTACAA	300
Db	320	CTGCCGGTTGATTTCTTTTGTCTCACCTGCGGTATGGTTACTGGGTCCTCGTGATTACAA	379
QY	301	TACACTACTTTTGGCAATGGGGGGAGATTAATTGATTAACAGACTTCAAGCTTTGAGCC	360
Db	380	TACACTACTTTTGGCAATGGGGGGAGATTAATTGATTAACAGACTTCAAGCTTTGAGCC	439
QY	361	CGGATTTTCTATGACACTGACATGACATGACATGCTGTGTAGTTTGAACCTTGTGTGTAG	420
Db	440	CGGATTTTCTATGACACTGACATGACATGACATGCTGTGTAGTTTGAACCTTGTGTGTAG	499
QY	421	CCGTGATTTGCTGACTCTTGCCAGCCCTCAGAAAGCATTTTAAAGTCAAGCAGAGACAA	480
Db	500	CCGTGATTTGCTGACTCTTGCCAGCCCTCAGAAAGCATTTTAAAGTCAAGCAGAGACAA	559
QY	481	GAGAGAGTTAGTGCGCAGCTCCCGGTGGACATCACCTGATCCCTTGAAGACAGACTTGTG	540
Db	560	GAGAGAGTTAGTGCGCAGCTCCCGGTGGACATCACCTGATCCCTTGAAGACAGACTTGTG	619
QY	541	AAGTCAGAGCTGTACACATTTGAATCTCAAGTCGAGCTTGTAGATTCGATGATTACAGA	600
Db	620	AAGTCAGAGCTGTACACATTTGAATCTCAAGTCGAGCTTGTAGATTCGATGATTACAGA	679
QY	601	AGAAAGATTTCTGAGTTTGGAAAGCAAAATGCACTGAACAGCAACCAATCCAAATGTTGTA	660
Db	680	AGAAAGATTTCTGAGTTTGGAAAGCAAAATGCACTGAACAGCAACCAATCCAAATGTTGTA	739
QY	661	ATTGAAGACTTTGAGTCTCACTTACCCGTTGCGTACCCCACTCTCACAAAGCTCTCTG	720
Db	740	ATTGAAGACTTTGAGTCTCACTTACCCGTTGCGTACCCCACTCTCACAAAGCTCTCTG	799
QY	721	AATATTTCTGTGTTTACCCCGAAGATATTTACAGTACATCTGCAGAGTCTCTTGGCCAG	780
Db	800	AATATTTCTGTGTTTACCCCGAAGATATTTACAGTACATCTGCAGAGTCTCTTGGCCAG	859
QY	781	GAGGAAGCCAGATCTGTGACTTGCAGAGATTCAGATTTTTCAGATGCCAATTTTCAAG	840
Db	860	GAGGAAGCCAGATCTGTGACTTGCAGAGATTCAGATTTTTCAGATGCCAATTTTCAAG	919
QY	841	GCAAGTTCAACTTACGAAATGATGCCATATAAAACACTCTGCTGTAGATTGGACATT	900
Db	920	GCAAGTTCAACTTACGAAATGATGCCATATAAAACACTCTGCTGTAGATTGGACATT	979
QY	901	TCAATATCAGACTTTTCTATCAGCTCGAGATGCTTTCAGCTGATCTGCCCTTAACGT	960
Db	980	TCAATATCAGACTTTTCTATCAGCTCGAGATGCTTTCAGCTGATCTGCCCTTAACGT	1039
QY	961	GATTCTAGAGTAACAAAGCTTACTCCAAAGACTGACGCTTGAAGATTAAGAGACACTGC	1020
Db	1040	GATTCTAGAGTAACAAAGCTTACTCCAAAGACTGACGCTTGAAGATTAAGAGACACTGC	1099
QY	1021	GTCCTTTTGAATAATAAAGCAGACACAAAGAGAAAGAGCTACTTACCACATATA	1080
Db	1100	GTCCTTTTGAATAATAAAGCAGACACAAAGAGAAAGAGCTACTTACCACATATA	1159
QY	1081	CTGCGGGAGTCTCTCCAGTTCAATTTTAACTGTGTCTGTGAATCCGAGCAATTCCT	1140
Db	1160	CTGCGGGAGTCTCTCCAGTTCAATTTTAACTGTGTCTGTGAATCCGAGCAATTCCT	1219

1141 AAAAGGCGATTTTGGAGCCCTTGTGACTATACAGTGAAGTGTGAAAAAGCGCAG 1200  
1220 AAAAGGCGATTTTGGAGCCCTTGTGACTATACAGTGAAGTGTGAAAAAGCGCAG 1279  
1201 CTAAGAGAGCTGTGCTAAACAGAGGGGCGCCGATTAAGCCGCTTTGTAGAGATGCC 1260  
1280 CTAAGAGAGCTGTGCTAAACAGAGGGGCGCCGATTAAGCCGCTTTGTAGAGATGCC 1339  
1261 TGGGCGCTGTGTGGATCTCCGCGCTTCCCTTGTGGCAGCAGCACTCACTCTC 1320  
1340 TGGGCGCTGTGTGGATCTCCGCGCTTCCCTTGTGGCAGCAGCACTCACTCTC 1399  
1321 CTGCTGGAACATCTTCTAACTTCAACCCAGACATATTCGTGTGCAAGCTCAAGTTTA 1380  
1400 CTGCTGGAACATCTTCTAACTTCAACCCAGACATATTCGTGTGCAAGCTCAAGTTTA 1459  
1381 TTTCACCCAGAAAGCTCCATTTTGTCTTCAACATTTGTGAATTTCTGTCTACTGCCACA 1440  
1460 TTTCACCCAGAAAGCTCCATTTTGTCTTCAACATTTGTGAATTTCTGTCTACTGCCACA 1519  
1441 AAGAGGCTCTGCGGAGGAGATATGATAGAGCTGGCTGGCTTGTGTGCTTCAATT 1500  
1520 AAGAGGCTCTGCGGAGGAGATATGATAGAGCTGGCTGGCTTGTGTGCTTCAATT 1579  
1501 CTTCAGCCCAACATACATGATCCATGAGAGCAGCGGAAAGCCCTGGCTCTTAAGATA 1560  
1580 CTTCAGCCCAACATACATGATCCATGAGAGCAGCGGAAAGCCCTGGCTCTTAAGATA 1639  
1561 TCCATCTCTCTCTGAAACAAATTTCTTTCACCTTACAGATGACCCCTCAATCCCATC 1620  
1640 TCCATCTCTCTCTGAAACAAATTTCTTTCACCTTACAGATGACCCCTCAATCCCATC 1699  
1621 ATATAGTGTGGTCCAGGAAACCGGCGATAGCCGCTTATGGGTTCTTACAACTAGAGAG 1680  
1700 ATATAGTGTGGTCCAGGAAACCGGCGATAGCCGCTTATGGGTTCTTACAACTAGAGAG 1759  
1681 AAATCTCAAGAAACAACCCAGATGAGAAATTTTGGAGCAATGTGTGTTTTTGGCTGC 1740  
1760 AAATCTCAAGAAACAACCCAGATGAGAAATTTTGGAGCAATGTGTGTTTTTGGCTGC 1819  
1741 AAGCATAGAGATAGGATTAATCTATTCAAGAAAGCTCAACATTTCTTAAAGCATGGG 1800  
1820 AAGCATAGAGATAGGATTAATCTATTCAAGAAAGCTCAACATTTCTTAAAGCATGGG 1879  
1801 ACTTAACTCATCTAAAGGTTTCTCTCAAGAGATGCTCTCTTGGGAGAGAGAAAGCC 1860  
1880 ACTTAACTCATCTAAAGGTTTCTCTCAAGAGATGCTCTCTTGGGAGAGAGAAAGCC 1939  
1861 CCAGCAAGATATGTAACAAGCAATCCAGCTTCATGGCCAGAGAGGTGGCGAAGATCTTC 1920  
1940 CCAGCAAGATATGTAACAAGCAATCCAGCTTCATGGCCAGAGAGGTGGCGAAGATCTTC 1999  
1921 CTTCAGAGAAACCGCATATTTATGTGTGTGAGATGCAAGAAATATGCGCAAGATGTA 1980  
2000 CTTCAGAGAAACCGCATATTTATGTGTGTGAGATGCAAGAAATATGCGCAAGATGTA 2059  
1981 CATTAGTCCCTTGTGCAATTAATTAACCAAGAGCTTGAAGAAATTAAGCAAGCAATG 2040  
2060 CATTAGTCCCTTGTGCAATTAATTAACCAAGAGCTTGAAGAAATTAAGCAAGCAATG 2119  
2041 AAAACCTGGCGCACTTTAAAGAAAGAAACGCTACCTTCAGAGATATTTGGTCAATA 2097  
2120 AAAACCTGGCGCACTTTAAAGAAAGAAACGCTACCTTCAGAGATATTTGGTCAATA 2176

RESULT 4  
AF121214 3291 bp mRNA linear PRI 12-DEC-1999  
LOCUS AF121214 Homo sapiens methionine synthase reductase (MTRR) mRNA, complete  
DEFINITION cds.  
ACCESSION AF121214  
VERSION AF121214.1 GI:6561338  
KEYWORDS

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
AUTHORS  
1 (bases 1 to 3291)  
Leclier, D., Odievre, M., Wu, Q., Wilson, A., Huizenga, J. U., Rozen, R.,  
Scherer, S. W., and Gravel, R. A.  
TITLE  
Molecular cloning, expression and physical mapping of the human  
methionine synthase reductase gene  
JOURNAL  
Gene 240 (1), 75-88 (1999)  
MEDLINE  
20033550  
PUBMED  
10564814  
REFERENCE  
2 (bases 1 to 3291)  
Leclier, D., Odievre, M., Wu, Q., Wilson, A., Huizenga, J. U.,  
Johns, T., Shoubiridge, E. A., Rosenblatt, D. S., Scherer, S. W., Rozen, R.,  
and Gravel, R. A.  
TITLE  
Direct Submision  
JOURNAL  
Submitted (18-JAN-1999) Human Genetics, Montreal Children's  
Hospital, 1460 Ste-Catherine west, Montreal, Quebec H3Z 2Z3, Canada  
FEATURES  
source  
1..3291  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/map="5p15.2-p15.3; between W11755 and D5S1957"  
1..3291  
/gene="MTRR"  
1..30  
/gene="MTRR"  
/evidence=not\_experimental  
31..2208  
/gene="MTRR"  
/EC\_number="2.1.1.135"  
/function="maintenance of methionine synthase in an active  
state"  
/note="alternatively spliced; alternative splicing results  
in extended 5' sequence compared to GenBank Accession  
Number AF025794; B12 vitamin; cblE complementation group"  
/codon\_start=1  
/product="methionine synthase reductase"  
/protein\_id="AA16876.1"  
/db\_xref="GI:6561339"  
/translation="MGASVRAAGARLVAVLCSFTVTCLEVRERFLLYATOGGAKA  
IAEMCEQAVVHGSPADLHCTSESRYDITETAPLVVVSTTGGDPDPTARKFVKE  
IONQTPDPPFPAHLKYGGLGSDSEYTYCNGSKIIIDKLGELGARHPDITGADVCV  
GLELVEEPIAGLWPALEKHFPSRSKQERISGALPVPASPLKTDLVKSELHTBSOV  
ELIARPDGSGRDKSEVLKQNAVNSQNSNVVIEDFESSLTRSVPLSQASLNIIPDLPEY  
LOVHLOESLQGESQSVTSADPQVPIKAVQLTNDALITTLTLELIDISTDPSY  
QPGDAFSVLCPSNDSVEVSLQRLQLEDKREHCVLTKIADTRKKGATLPOHLPACS  
LOFITWLCLEIRAIPIKAFIRLALVDYTSDAKRLQELCSKQGAADVSFPYDAAC  
LIDLILAPSCQPLSLILEHLPIKQPRYSQASSLPHRGCLHPVFNIVFELSTATT  
EVLKRGVCTGWLALVAVLYLPNIHSHSDSKALAPKISIPRTNSFHLPPDSIP  
ILMVPGTGIAPIFGLQREKLQHPDNGGAMWLPFGCHHKORDYFRLRLRLFL  
KHGIIHTLKVSRSDAPVGESEBAPAKYVDNIQLHGQVARTLLQENGHIYYCGDAKN  
MAKQVHDALVQIISKEVGKLEAMKTLATLKEKRYLQDIMS"  
transit\_peptide 31..123  
/gene="MTRR"  
/evidence=not\_experimental  
3' UTR  
2209..3291  
/gene="MTRR"  
/evidence=not\_experimental  
polyA\_signal  
3246..3251  
/gene="MTRR"  
/evidence=not\_experimental  
polyA\_site  
3268  
/gene="MTRR"  
/evidence=experimental  
polyA\_site  
3276  
/gene="MTRR"  
/evidence=experimental  
ORIGIN

Query Match 100.0%; Score 2097; DB 9; Length 3291;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATAGAGGGTTCTGTTACTATATGCTACACAGAGGACAGGAAAGCCATCCGCAAA 60  
DB 112 ATAGAGGGTTCTGTTACTATATGCTACACAGAGGACAGGAAAGCCATCCGCAAA 171  
QY 61 GAATGTGTGAGCAAGCTGTGTACATGATTTCTGCAGATCTTCACTGTATTAAGTAA 120  
DB 172 GAATGTGTGAGCAAGCTGTGTACATGATTTCTGCAGATCTTCACTGTATTAAGTAA 231  
QY 121 TCCGATTAAGTATGACTTAACAAACCGAAACAGCTCTCTGTGTGTGTGTCTTACCA 180  
DB 232 TCCGATTAAGTATGACTTAACAAACCGAAACAGCTCTCTGTGTGTGTGTCTTACCA 291  
QY 181 GGACCGGAGACCAACCGACACAGCCCGCAAGTTTGTAAAGAAATACGAACCAACA 240  
DB 292 GGACCGGAGACCAACCGACACAGCCCGCAAGTTTGTAAAGAAATACGAACCAACA 351  
QY 241 CTGCGGTTGATTTCTTTGCTCACTGCGGTATGGTTACTGGGTCTCGGTATTCAGAA 300  
DB 352 CTGCGGTTGATTTCTTTGCTCACTGCGGTATGGTTACTGGGTCTCGGTATTCAGAA 411  
QY 301 TACACCTACTTTGCAATGGGGGGAAGATTAATGATTAAGCACTTCAAGCTTGAACC 360  
DB 412 TACACCTACTTTGCAATGGGGGGAAGATTAATGATTAAGCACTTCAAGCTTGAACC 471  
QY 361 CGGATTTCTATGACACTGACATGACAGATGACGTGTAGTTTGAACCTTGTGTGAG 420  
DB 472 CGGATTTCTATGACACTGACATGACAGATGACGTGTGTAGTGAACCTTGTGTGAG 531  
QY 421 CCGTGATTTGCTGACCTGCGCAGCCCTCAGAAAGATTTTAAAGTCAAGCAGAGCAAA 480  
DB 532 CCGTGATTTGCTGACCTGCGCAGCCCTCAGAAAGATTTTAAAGTCAAGCAGAGCAAA 591  
QY 481 GAGGAGTAAGTGGCGCACTCCCGGTGACATCCTGATCTTGAAGACAGACCTTGTG 540  
DB 592 GAGGAGTAAGTGGCGCACTCCCGGTGACATCCTGATCTTGAAGACAGACCTTGTG 651  
QY 541 AAGTCAGAGCTGTACACTGAATTCGAAGTTCGAGCTTGCAGATTCGATTTACGAGA 600  
DB 652 AAGTCAGAGCTGTACACTGAATTCGAAGTTCGAGCTTGCAGATTCGATTTACGAGA 711  
QY 601 AGAAAGATTCGAGGTTTGAACAAATGCAAGTGAACAGCAACCAATCCATGTTGTA 660  
DB 712 AGAAAGATTCGAGGTTTGAACAAATGCAAGTGAACAGCAACCAATCCATGTTGTA 771  
QY 661 ATTGAAGCTTTGAGTCTCACTTACCCGTTGGTACCCCACTCTCAAGCCCTCTCTG 720  
DB 772 ATTGAAGCTTTGAGTCTCACTTACCCGTTGGTACCCCACTCTCTCAAGCCCTCTCTG 831  
QY 721 AATATTCCTGTTTACCCCGAATATTTAAGGATCACTGAGAGTCTCTTGGGCGAG 780  
DB 832 AATATTCCTGTTTACCCCGAATATTTAAGGATCACTGAGAGTCTCTTGGGCGAG 891  
QY 781 GAGGAAAGCCAAAGTATCTGTGACTTCAGCAGATCCAGTTTTCAGTGCCAAATTCAA 840  
DB 892 GAGGAAAGCCAAAGTATCTGTGACTTCAGCAGATCCAGTTTTCAGTGCCAAATTCAA 951  
QY 841 GCAATTCACTTACGAAATGATGCCATTAACCACTGCTGTGTGAATTTGACATT 900  
DB 952 GCAATTCACTTACGAAATGATGCCATTAACCACTGCTGTGTGAATTTGACATT 1011  
QY 901 TCAATTCAGACTTTCTTATCAGCTGAGATGCTTACGCGATCTGCGCTTAACAGT 960  
DB 1012 TCAATTCAGACTTTCTTATCAGCTGAGATGCTTACGCGATCTGCGCTTAACAGT 1071  
QY 961 GATTCTGAGGTACAAAGCCTACTCCAAAGCTGACGCTTGAAGTAAAGAGACACTGC 1020  
DB 1072 GATTCTGAGGTACAAAGCCTACTCCAAAGCTGACGCTTGAAGTAAAGAGACACTGC 1131  
QY 1021 GTCTTTTGAATTAAGGACAGACAAAGAAAGAGCTACCTTACCCAGCATATA 1080

DB 1132 GTCTTTTGAATTAAGGACAGACAAAGAAAGAGCTTACCTTACCCAGCATATA 1191  
QY 1081 CCGCGGAGATGTTCTCTCCAGATTCATTTTAACTGGGTCTTGAATCCGAGCAATTCCT 1140  
DB 1192 CCGCGGAGATGTTCTCTCCAGATTCATTTTAACTGGGTCTTGAATCCGAGCAATTCCT 1251  
QY 1141 AAAAGCAATTTTTCGAGCCCTTGTGACTATACAGTACAGTCTGAAAAAGCCGAGG 1200  
DB 1252 AAAAGCAATTTTTCGAGCCCTTGTGACTATACAGTACAGTCTGAAAAAGCCGAGG 1311  
QY 1201 CTACAGAGCTGTGCAAGTAAACAAAGGAGCCGATTAATAGCCGCTTGTAGAGATGCC 1260  
DB 1312 CTACAGAGCTGTGCAAGTAAACAAAGGAGCCGATTAATAGCCGCTTGTAGAGATGCC 1371  
QY 1261 TGTGCTGCTGTTGTGATCTCTCTGCTTCCCTTCTTCCAGCCACCACTCAGTCTC 1320  
DB 1372 TGTGCTGCTGTTGTGATCTCTCTCTGCTTCCCTTCTTCCAGCCACCACTCAGTCTC 1431  
QY 1321 CTGCTGAACATCTTCTTAACTTCAACCCAGACATATTCGTGCAAGCTCAAGTTTA 1380  
DB 1432 CTGCTGAACATCTTCTTAACTTCAACCCAGACATATTCGTGCAAGCTCAAGTTTA 1491  
QY 1381 TTTCAACCGAGAAAGCTCCATTTTGTCTTCAATTTGTGAATTTCTGTCTACTGCCACA 1440  
DB 1492 TTTCAACCGAGAAAGCTCCATTTTGTCTTCAATTTGTGAATTTCTGTCTACTGCCACA 1551  
QY 1441 ACAGAGTTCTGCGGAAAGGAGATATGATACAGCTGCGTGGCTTGTGTGCTCAGTT 1500  
DB 1552 ACAGAGTTCTGCGGAAAGGAGATATGATACAGCTGCGTGGCTTGTGTGCTCAGTT 1611  
QY 1501 CTTCAAGCAAACTACATGATCATCCATGAAGACAGCGGAAAGCCCTGCTCTTAAGATA 1560  
DB 1612 CTTCAAGCAAACTACATGATCATCCATGAAGACAGCGGAAAGCCCTGCTCTTAAGATA 1671  
QY 1561 TCCATCTCTCTGGAACAAATTTCTTCCACTTACAGATACACCCCTCAATCCCATC 1620  
DB 1672 TCCATCTCTCTGGAACAAATTTCTTCCACTTACAGATACACCCCTCAATCCCATC 1731  
QY 1621 ATATAGTGGGTCAAGAAACCGGCAATAGCCCGTTTATTTGGTCTTACAAATAGAGAG 1680  
DB 1732 ATATAGTGGGTCAAGAAACCGGCAATAGCCCGCGTTTATTTGGTCTTACAAATAGAGAG 1791  
QY 1681 AAATCTCAAGAACAAACCCAGATGAAATTTTGAAGCAATGTGTTTGTGCTGC 1740  
DB 1792 AAATCTCAAGAACAAACCCAGATGAAATTTTGAAGCAATGTGTTTGTGCTGC 1851  
QY 1741 AGGCAATAGATAGGATTTATCTTTCAGAAAGAGCTCAGACATTTCTTAAGCATGGG 1800  
DB 1852 AGGCAATAGATAGGATTTATCTTTCAGAAAGAGCTCAGACATTTCTTAAGCATGGG 1911  
QY 1801 ATCTTAATCATCTAAAGGTTTCTCTCAAGAGATGCTCTGTGGGAGAGAGAGGCC 1860  
DB 1912 ATCTTAATCATCTAAAGGTTTCTCTCTCAAGAGATGCTCTGTGGGAGAGAGAGGCC 1971  
QY 1861 CAGCAAAATATATACAAAGCAATTCAGCTTCAATGCGCAGAGAGTGGCGAATCTTC 1920  
DB 1972 CAGCAAAATATATACAAAGCAATTCAGCTTCAATGCGCAGAGAGTGGCGAATCTTC 2031  
QY 1921 CTCGAGAGAAAGGCAATTTATTTGTGTGTGAGATGCAAGAAATATGCGCAAGATGTA 1980  
DB 2032 CTCGAGAGAAAGGCAATTTATTTGTGTGTGAGATGCAAGAAATATGCGCAAGATGTA 2091  
QY 1981 CATGATCCCTTGTGCAATATATAGCAAGAGGTTGAGTTGAAAACTAGAAAGCAATG 2040  
DB 2092 CATGATCCCTTGTGCAATATATAGCAAGAGGTTGAGTTGAAAACTAGAAAGCAATG 2151  
QY 2041 AAAACCTGCGCACTTTAAAGAAAGAAAAAGCTACCTTCAAGATATTTGTCATTA 2097  
DB 2152 AAAACCTGCGCACTTTAAAGAAAGAAAAAGCTACCTTCAAGATATTTGTCATTA 2208

RESULT 5



LOCUS CO726091 3241 bp DNA linear PAT 03-FEB-2004  
DEFINITION Sequence 12025 from Patent WO02068579.  
ACCESSION CO726091  
VERSION CO726091.1 GI:42288134  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
Kites, such as nucleic acid arrays, comprising a majority of  
humaneons or transcripts, for detecting expression and other uses  
thereof  
Patent: WO 02068579-A 12025 06-SEP-2002;  
JOURNAL  
PE Corporation  
FEATURES  
source 1..3241  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 95.1%; Score 1995; DB 6; Length 3241;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 205; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ATGAGAGGTTCTGTACTATATGCTACACAGCAGGACAGGCAAAAGCCATCGAGAA 60  
DB ATGAGAGGTTCTGTACTATATGCTACACAGCAGGACAGGCAAAAGCCATCGAGAA 139  
QY 61 GAAATGTGAGCAGAGCTGTGTACATGATTTTCTGCAGATCTTCACTGATTAAGTAA 120  
DB GAAATGTGAGCAGAGCTGTGTACATGATTTTCTGCAGATCTTCACTGATTAAGTAA 199  
QY 121 TCCGATTAAGTACCTAAACCGAAACAGCTCTTGTGTGTGTTCTACACG 180  
DB TCCGATTAAGTACCTAAACCGAAACAGCTCTTGTGTGTGTTCTACACG 259  
QY 200 TCCGATTAAGTACCTAAACCGAAACAGCTCTTGTGTGTGTTCTACACG 259  
DB TCCGATTAAGTACCTAAACCGAAACAGCTCTTGTGTGTGTTCTACACG 259  
QY 181 GGCACCGGAGACCCACCCGACAGAGCCGCAAGTTGTTAAGAAATACAGAACAA 240  
DB GGCACCGGAGACCCACCCGACAGAGCCGCAAGTTGTTAAGAAATACAGAACAA 319  
QY 241 CTGCCGCTGATTTCTTGTCTCACTGCGGATGAGTTACTGGGCTCGGATTCAGAA 300  
DB CTGCCGCTGATTTCTTGTCTCACTGCGGATGAGTTACTGGGCTCGGATTCAGAA 379  
QY 320 CTGCCGCTGATTTCTTGTCTCACTGCGGATGAGTTACTGGGCTCGGATTCAGAA 379  
DB CTGCCGCTGATTTCTTGTCTCACTGCGGATGAGTTACTGGGCTCGGATTCAGAA 379  
QY 301 TACACCTACTTTGCAATGGGGGGAAGATTAATTGATTAACGACTTGAAGGCTTGAAGC 360  
DB TACACCTACTTTGCAATGGGGGGAAGATTAATTGATTAACGACTTGAAGGCTTGAAGC 439  
QY 380 TACACCTACTTTGCAATGGGGGGAAGATTAATTGATTAACGACTTGAAGGCTTGAAGC 439  
DB TACACCTACTTTGCAATGGGGGGAAGATTAATTGATTAACGACTTGAAGGCTTGAAGC 439  
QY 361 CGGCACTTTCTATGACACTGACATGAGATGACTGTGATGTTAGAACTTGTGTGAG 420  
DB CGGCACTTTCTATGACACTGACATGAGATGACTGTGATGTTAGAACTTGTGTGAG 499  
QY 440 CGGCACTTTCTATGACACTGACATGAGATGACTGTGATGTTAGAACTTGTGTGAG 499  
DB CGGCACTTTCTATGACACTGACATGAGATGACTGTGATGTTAGAACTTGTGTGAG 499  
QY 421 CGGTGATGCTGAGCTCTGCGCAAGCCCTTCAAGAAAGATTTAGTCAAGCAGAGCAA 480  
DB CGGTGATGCTGAGCTCTGCGCAAGCCCTTCAAGAAAGATTTAGTCAAGCAGAGCAA 559  
QY 500 CGGTGATGCTGAGCTCTGCGCAAGCCCTTCAAGAAAGATTTAGTCAAGCAGAGCAA 559  
DB CGGTGATGCTGAGCTCTGCGCAAGCCCTTCAAGAAAGATTTAGTCAAGCAGAGCAA 559  
QY 481 GAGGAATAGTGGCCCACTCCCGGTGGATTCACCTGCATCCCTTGAAGGACAGACTTGG 540  
DB GAGGAATAGTGGCCCACTCCCGGTGGATTCACCTGCATCCCTTGAAGGACAGACTTGG 619  
QY 560 GAGGAATAGTGGCCCACTCCCGGTGGATTCACCTGCATCCCTTGAAGGACAGACTTGG 619  
DB GAGGAATAGTGGCCCACTCCCGGTGGATTCACCTGCATCCCTTGAAGGACAGACTTGG 619  
QY 541 AAGTCAGAGCTGTACACATTTGATCTCAAGTCGAGCTTCTGATTCGATTCGATTCAGGA 600  
DB AAGTCAGAGCTGTACACATTTGATCTCAAGTCGAGCTTCTGATTCGATTCGATTCAGGA 679  
QY 620 AAGTCAGAGCTGTACACATTTGATCTCAAGTCGAGCTTCTGATTCGATTCGATTCAGGA 679  
DB AAGTCAGAGCTGTACACATTTGATCTCAAGTCGAGCTTCTGATTCGATTCGATTCAGGA 679  
QY 601 AGAAGAGATTCGAGCTTTGAAGCAAAATGACAGTGAACAGCAACCAATCAATGTTGTA 660  
DB AGAAGAGATTCGAGCTTTGAAGCAAAATGACAGTGAACAGCAACCAATCAATGTTGTA 739  
QY 680 AGAAGAGATTCGAGCTTTGAAGCAAAATGACAGTGAACAGCAACCAATCAATGTTGTA 739  
DB AGAAGAGATTCGAGCTTTGAAGCAAAATGACAGTGAACAGCAACCAATCAATGTTGTA 739  
QY 661 ATTGAAGACTTTGAGTCTCACTTACCCGTTGGGTACCCCACTCTCAAGAGCTCTCTG 720  
DB ATTGAAGACTTTGAGTCTCACTTACCCGTTGGGTACCCCACTCTCAAGAGCTCTCTG 799

DB ATTGAAGACTTTGAGTCTCACTTACCCGTTGGGTACCCCACTCTCAAGAGCTCTCTG 799  
QY 721 AATATTCCTGCTTACCCCAAGATATTTACAGGTAATCTGACAGAGCTCTTGGCCAG 780  
DB AATATTCCTGCTTACCCCAAGATATTTACAGGTAATCTGACAGAGCTCTTGGCCAG 859  
QY 781 GAGGAAGCAGATCTGCTGACTTCAAGATTCGAGATTCGAGATTCGAGATTCGAGATTCGAG 840  
DB GAGGAAGCAGATCTGCTGACTTCAAGATTCGAGATTCGAGATTCGAGATTCGAGATTCGAG 919  
QY 841 GCAATTCACCTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
DB GCAATTCACCTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 979  
QY 920 GCAATTCACCTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 979  
DB GCAATTCACCTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 979  
QY 901 TCAATTAAGACTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 960  
DB TCAATTAAGACTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1039  
QY 980 TCAATTAAGACTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1039  
DB TCAATTAAGACTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1039  
QY 961 GATTCGAGTACAAAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1020  
DB GATTCGAGTACAAAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1099  
QY 1040 GATTCGAGTACAAAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1099  
DB GATTCGAGTACAAAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1099  
QY 1021 GTCTTTGAAATTAAGCAGACACAAAGAAAGAGCTTACCTTACCCGACATATA 1080  
DB GTCTTTGAAATTAAGCAGACACAAAGAAAGAGCTTACCTTACCCGACATATA 1159  
QY 1100 GTCTTTGAAATTAAGCAGACACAAAGAAAGAGCTTACCTTACCCGACATATA 1159  
DB GTCTTTGAAATTAAGCAGACACAAAGAAAGAGCTTACCTTACCCGACATATA 1159  
QY 1081 CCTGGGGAGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1140  
DB CCTGGGGAGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1219  
QY 1160 CCTGGGGAGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1219  
DB CCTGGGGAGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1219  
QY 1141 AAAAGGCAATTTTGGAGCCCTTGTGATTAACAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAG 1200  
DB AAAAGGCAATTTTGGAGCCCTTGTGATTAACAGAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAG 1279  
QY 1220 AAAAGGCAATTTTGGAGCCCTTGTGATTAACAGAGAGAGCTGAGAGAGAGAGAGAGAGAGAG 1279  
DB AAAAGGCAATTTTGGAGCCCTTGTGATTAACAGAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAG 1279  
QY 1201 CTACAGAGCTGTGAGTAAACAG 1260  
DB CTACAGAGCTGTGAGTAAACAG 1339  
QY 1280 CTACAGAGCTGTGAGTAAACAG 1339  
DB CTACAGAGCTGTGAGTAAACAG 1339  
QY 1261 TGTGCTGCTTGTGATCT 1320  
DB TGTGCTGCTTGTGATCT 1399  
QY 1340 TGTGCTGCTTGTGATCT 1399  
DB TGTGCTGCTTGTGATCT 1399  
QY 1321 CTGCTGGAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1380  
DB CTGCTGGAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1459  
QY 1400 CTGCTGGAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1459  
DB CTGCTGGAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1459  
QY 1381 TTTTCAACCGAGAAAGCTCAATTTTCTTCAATTTGGAATTTCTGCTACTGACACA 1440  
DB TTTTCAACCGAGAAAGCTCAATTTTCTTCAATTTGGAATTTCTGCTACTGACACA 1519  
QY 1460 TTTTCAACCGAGAAAGCTCAATTTTCTTCAATTTGGAATTTCTGCTACTGACACA 1519  
DB TTTTCAACCGAGAAAGCTCAATTTTCTTCAATTTGGAATTTCTGCTACTGACACA 1519  
QY 1441 ACAGAGTTCTGCGAGAGAGAGATGATACAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1500  
DB ACAGAGTTCTGCGAGAGAGAGATGATACAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1579  
QY 1520 ACAGAGTTCTGCGAGAGAGAGATGATACAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1579  
DB ACAGAGTTCTGCGAGAGAGAGATGATACAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1579  
QY 1501 CTTCAGCCGAAACATACATGATCCCATGAAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560  
DB CTTCAGCCGAAACATACATGATCCCATGAAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1639  
QY 1580 CTTCAGCCGAAACATACATGATCCCATGAAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1639  
DB CTTCAGCCGAAACATACATGATCCCATGAAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1639  
QY 1561 TCCATCTCTCTGAAACAAATTTCTTCAATTTGGAATTTCTGCTACTGACACA 1620  
DB TCCATCTCTCTGAAACAAATTTCTTCAATTTGGAATTTCTGCTACTGACACA 1699  
QY 1640 TCCATCTCTCTGAAACAAATTTCTTCAATTTGGAATTTCTGCTACTGACACA 1699  
DB TCCATCTCTCTGAAACAAATTTCTTCAATTTGGAATTTCTGCTACTGACACA 1699  
QY 1621 ATTAATGAGGCTTCAAGAAACAGGATAGCCCGGTTATTTGAGTTCTTACAAATAGAGAG 1680  
DB ATTAATGAGGCTTCAAGAAACAGGATAGCCCGGTTATTTGAGTTCTTACAAATAGAGAG 1759  
QY 1700 ATTAATGAGGCTTCAAGAAACAGGATAGCCCGGTTATTTGAGTTCTTACAAATAGAGAG 1759  
DB ATTAATGAGGCTTCAAGAAACAGGATAGCCCGGTTATTTGAGTTCTTACAAATAGAGAG 1759  
QY 1681 AAATCTCAAGAAACAAACCAATGGAATTTTGAAGCAATGAGGATTTTGTGCTGCTG 1740  
DB AAATCTCAAGAAACAAACCAATGGAATTTTGAAGCAATGAGGATTTTGTGCTGCTG 1819  
QY 1760 AAATCTCAAGAAACAAACCAATGGAATTTTGAAGCAATGAGGATTTTGTGCTGCTG 1819  
DB AAATCTCAAGAAACAAACCAATGGAATTTTGAAGCAATGAGGATTTTGTGCTGCTG 1819  
QY 1741 AGGATTAAGATGAGGATTTATCTTCAAGAAAGAGCTCAGACATTTCTTAAAGATGAG 1800  
DB AGGATTAAGATGAGGATTTATCTTCAAGAAAGAGCTCAGACATTTCTTAAAGATGAG 1800

Db 1820 AGGCATAGAGTATGAGTATTAATTCAGAAAAAGAGCTCAGACATTTCTTAAGCATGG 1879

Qy 1801 ATCTTAATCATCTTAAGAGTTTCCTTCTCAAGAGATGCTCTGTTGGGAGAGAGAGCC 1860

Db 1880 ATCTTAATCATCTTAAGAGTTTCCTTCTCAAGAGATGCTCTGTTGGGAGAGAGAGCC 1939

Qy 1861 CCAGCAAGATATGTACAAGACATCAGCTTCATGCGCAGCAGGTGGCGAGATCTTC 1920

Db 1940 CCAGCAAGATATGTACAAGACATCAGCTTCATGCGCAGCAGGTGGCGAGATCTTC 1999

Qy 1921 CTCGAGAGAGAGCCATATTTATGTTGTGTGTGAGATGCAAGATATGCGCCAGAGATGA 1980

Db 2000 CTCGAGAGAGAGCCATATTTATGTTGTGTGTGAGATGCAAGATATGCGCCAGAGATGA 2059

Qy 1981 CATGATCCCTTGTGCAATATATATAGCAAGAGAGGTGTGAGTGAATAAGTAAAGCATG 2040

Db 2060 CATGATCCCTTGTGCAATATATATAGCAAGAGAGGTGTGAGTGAATAAGTAAAGCATG 2119

Qy 2041 AAAACCTGGCCACTTTAAAGAGAAAAAGCTACCTTCAGATATTTGGTCATAA 2097

Db 2120 AAAACCTGGCCACTTTAAAGAGAAAAAGCTACCTTCAGATATTTGGTCATAA 2176

RESULT 6  
BC054816  
LOCUS  
DEFINITION  
Accession  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS

BC054816 3310 bp. mRNA linear PRI 16-SEP-2003  
Homo sapiens 5-methyltetrahydrofolate-homocysteine  
methyltransferase reductase, mRNA (cdna clone IMAGE:5205285),  
partial cds.  
BC054816  
BC054816.1 GI:33392775

Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 3310)  
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,  
Altschul,S.F., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,  
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heide,F.,  
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,  
Scapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,  
Schaefer,T.E., Brownstein,M.J., Ueda,T.B., Toshiki,Y.K.I.,  
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,  
Abramson,R.D., Muliahy,S.J., Bosak,S.A., McEwan,P.J.,  
Mokernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,  
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.D., Hulyk,S.W.,  
Villalón,D.K., Wuzny,D.M., Sodergren,E.J., Lu,X., Gidde,R.A.,  
Fahney,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,  
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,  
Boiffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,  
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,  
Butterfield,Y.S., Krzywicki,M.I., Skalska,U., Smailus,D.E.,  
Scherer,A., Schein,J.E., Jones,S.J. and Marz,M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

2 (bases 1 to 3310)  
Strausberg,R.  
Direct Submission  
Submitted (03-JUL-2003) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed By: The I.M.A.G.E. Consortium (ULNL)

DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc\\_mgc@hgti.nih.gov](mailto:nisc_mgc@hgti.nih.gov)  
Ahter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,  
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,  
Diatchenko,L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,  
Hansen,N., Ho,S.-L., Karlén,E., Kwong,P., Latic,P., Legaspi,R.,  
Maduro,O.L., Masello,C., Maskeri,B., Mastrian,S.D., McLooney,J.C.,  
McDowell,J., Pearson,R., Stantrop,S., Thomas,P.J., Touchman,J.W.,  
Tsougen,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,  
Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/ULNL at: <http://image.llnl.gov>  
Series: IRAX Plate: 115 Row: d Column: 11  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 4505278.

FEATURES  
source  
1. .3310  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5205285"  
/issue\_type="Lung, Spleen, fetal, pooled"  
/clone\_lib="NIH MGC\_122"  
/lab\_host="DH10B"  
/note="Vector: pCMV-Sport6"  
<1. .3310  
/gene="WTRR"  
/note="synonym: MSR"  
/db\_xref="LocusID:4552"  
/db\_xref="MIM:602568"  
<1. 2149  
/gene="WTRR"  
/codon\_start=2  
/product="WTRR protein"  
/protein\_id="AAH54816.1"  
/db\_xref="GI:33392776"  
/db\_xref="LocusID:4552"  
/db\_xref="MIM:602568"  
/translation="VPARLARKEFTVTCLEWRRFLLYATQGAALAEICEQAV  
VHFSADLHICISDSRDYDLRTETAPLVVWSTGTDPPTPAKPKVEIQNTLPVF  
PALRYGLGLGDESEYVFGNGKIIIDKQELGAGHVFYVSTATTEVLRGVGTS  
AGMPALRKHFRRSGOEISGALVPASPSRSDLYKSLHIESOVLELRFDDSGR  
KQSEVYKQNAVNSQNSVPESSITVPSPLSQAISNIPLPREYOVHIOESIG  
QESQVTSVADVPFVYPSKAVQITNDATKTTLVLEIDISMTDSYOGDAFSYIC  
PNSDSVQSLQRLQEDKREKCVLTKIKADTKKATLRFQHIPIAGSCLOPITWCL  
IRAPPKAFKRALVDYTSDAEKRRLOELCSKQADYSFVBDACCLDLILAAPS  
CQPLPLLEHLPKLPDPYSCASSLFFHGKJHFNIVFVISTATTEVLRGVGTS  
WLLAVASVQPNIIHSHSDSGKALPKISISRTNSFLLPDPSPILPIVGPVGTI  
APRIGLOHREKIOHSDPENGAMLPFECRKHDDYLRKLRHPLKIGITLHKV  
SFRSDPVGEEERPAKYVDNIDHQQVARIILIOENGHLYVCGDAKMAKDVHDALV  
QITSKVGKELAMKTLATLKEKRILQDIMS  
59. .493  
/gene="WTRR"  
/note="Flavodoxin; Region: Flavodoxin"  
/db\_xref="CD:pfam00258"  
848. 2143  
/gene="WTRR"  
/note="CysJ; Region: Sulfite reductase, alpha subunit  
(flavoprotein) (Inorganic ion transport and metabolism)"  
/db\_xref="CD:COG0369"

ORIGIN  
Query Match 92.7%; Score 1944; DB 9; Length 3310;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2094; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGAGAGAGTTTCCTTACTATATGCTACACAGAGGAGAGGAAAGGCATGCGAGAA 60  
Db 53 ATGAGAGAGTTTCCTTACTATATGCTACACAGAGGAGAGGAAAGGCATGCGAGAA 112





SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
AUTHORS Nelson, R.M., Marnell, G., Kemmerer, S., Hoyal, C.R., Shi, M.M.,  
Cantor, C.R. and Braun, A.  
TITLE Large-Scale Validation of Single Nucleotide Polymorphisms in Gene  
Regions  
JOURNAL Genome Res. (2004) In press  
COMMENT Contact: Andreas Braun  
Pharmaceuticals division  
Sequenom, Inc.  
3595 John Hopkins Court, San Diego, CA 92121, USA  
Tel: 18582029018  
Fax: 18582029020  
Email: abraun@sequenom.com  
Primer A: No primer sequence submitted  
Primer B: No primer sequence submitted  
STS size: 2933.  
FEATURES  
source location/Qualifiers  
1..2933  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone\_id="Human DNA (Sequenom)"  
<1..>2933  
ORIGIN  
STS  
Query Match 63.8%; Score 1338; DB 11; Length 2933;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1778; Conservative 0; Mismatches 4; Indels 2; Gaps 2;  
QY 315 CAATGGGGGGAAGTAATGTGTAAGACCTCAAGAGCTGGAGCCCGCATTTCTATGA 374  
Db 2846 CAATGGGGGGAAGTAATGTGTAAGACCTCAAGAGCTGGAGCCCGCATTTCTATGA 2787  
QY 375 CACTGACATGACAGATGACTGTAGGTTTGAACCTTGTGGTTGAGCCGTGATTCCTGG 434  
Db 2786 CACTGACATGACAGATGACTGTAGGTTTGAACCTTGTGGTTGAGCCGTGATTCCTGG 2727  
QY 435 ACTTGGCCAGCCCTCAGAAAGCATTTTAAAGTCAAGCAGAGCAAGAGAGATTAAGTGG 494  
Db 2726 ACTTGGCCAGCCCTCAGAAAGCATTTTAAAGTCAAGCAGAGCAAGAGAGATTAAGTGG 2667  
QY 495 CGCAGTCCCGGTGACATCACCCTGATCCTTGAAGAGACCTTGAAGTCAAGAGTCT 554  
Db 2666 CGCAGTCCCGGTGACATCACCCTGATCCTTGAAGAGACCTTGAAGTCAAGAGTCT 2607  
QY 555 ACACATTGAATCTCAAGTCAAGCTTCTGAGATTCGATGATTCAGAGAAAGATTCGA 614  
Db 2606 ACACATTGAATCTCAAGTCAAGCTTCTGAGATTCGATGATTCAGAGAAAGATTCGA 2547  
QY 615 GGTTTGAAGCAAAATGAGTGAACAGCAACCAATTCATGTTGTAATTGAAGCTTTGA 674  
Db 2546 GGTTTGAAGCAAAATGAGTGAACAGCAACCAATTCATGTTGTAATTGAAGCTTTGA 2487  
QY 675 GTCTCACTTACCCGTTGGGTGATACCCCACTCTCAAGAGCTCTGATATATTCCTGGTT 734  
Db 2486 GTCTCACTTACCCGTTGGGTGATACCCCACTCTCAAGAGCTCTGATATATTCCTGGTT 2427  
QY 735 ACCCCAGAAATATTTACAGTATCTGACAGAGTCTTGGCCAGAGAGAAAGCAAGT 794  
Db 2426 ACCCCAGAAATATTTACAGTATCTGACAGAGTCTTGGCCAGAGAGAAAGCAAGT 2367  
QY 795 ATCTGTGACTTCAGAGATTCAGATTTTCAAGTCCCAATTTCAAGGAGATTCACTTAC 854  
Db 2366 ATCTGTGACTTCAGAGATTCAGATTTTCAAGTCCCAATTTCAAGGAGATTCACTTAC 2307  
QY 855 TACGATATGATGCATTAAGAACCACTGTGCTGATGAATTTGACATTTCAATTAAGACTT 914  
Db 2306 TACGATATGATGCATTAAGAACCACTGTGCTGATGAATTTGACATTTCAATTAAGACTT 2247

QY 915 TTCTATCAGCCCTGAGATGCTTTCAGCGTATCTGCCCTAACAGTATCTGAGGTACA 974  
Db 2246 TTCTATCAGCCCTGAGATGCTTTCAGCGTATCTGCCCTAACAGTATCTGAGGTACA 2187  
QY 975 AAGCTTATCTCAAGAGCTGAGCTTGAAGATTAAGAGAGCACTGGCTCTTTTGAAGAT 1034  
Db 2186 AAGCTTATCTCAAGAGCTGAGCTTGAAGATTAAGAGAGCACTGGCTCTTTTGAAGAT 2127  
QY 1035 AAGGAGACACAAAGAAAGAAAGAGCTACCTTACCCAGCATATACCTGCGGATTTTC 1094  
Db 2126 AAGGAGACACAAAGAAAGAAAGAGCTACCTTACCCAGCATATACCTGCGGATTTTC 2067  
QY 1095 TCTCAATTCATTTTACCTGCTGTGTGTAATTCGAGCAATTCCTTAAAGGCAATTTT 1154  
Db 2066 TCTCAATTCATTTTACCTGCTGTGTGTAATTCGAGCAATTCCTTAAAGGCAATTTT 2007  
QY 1155 GCGAGCCCTTGTGAGATTAATACAGTACAGTGTGTAAGAAAGCGAGCTACAGAGCTGTG 1214  
Db 2006 GCGAGCCCTTGTGAGATTAATACAGTACAGTGTGTAAGAAAGCGAGCTACAGAGCTGTG 1947  
QY 1215 CAGTAAACAAAGGGGAGCCGATTAATAGCCCTTGTACAGATGCTGTGCTTGT 1274  
Db 1946 CAGTAAACAAAGGGGAGCCGATTAATAGCCCTTGTACAGATGCTGTGCTTGT 1887  
QY 1275 GAGTCTCTCTCTGCTTCT 1334  
Db 1886 GAGTCTCTCTCTGCTTCT 1827  
QY 1335 TCTTAACTTCAACCCAGACATATTCGTGTGAGAGCTCAAGTTTATTCACAGAGAA 1394  
Db 1826 TCTTAACTTCAACCCAGACATATTCGTGTGAGAGCTCAAGTTTATTCACAGAGAA 1767  
QY 1395 GCTCCATTTTGTCTTCAACATTTGTGTGAAATTTCTGTACTGACCAACAGAGTCTGCG 1454  
Db 1766 GCTCCATTTTGTCTTCAACATTTGTGTGAAATTTCTGTACTGACCAACAGAGTCTGCG 1707  
QY 1455 GAAAGGAGTATGTACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1514  
Db 1706 GAAAGGAGTATGTACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1647  
QY 1515 ACATGATCCCATGAAAGACAGC--GGAAAGCCCTGCTCTTAAGATATCATCTCTC 1573  
Db 1646 ACATGATCCCATGAAAGACAGC--GGAAAGCCCTGCTCTTAAGATATCATCTCTC 1587  
QY 1574 GAACCAAAATTTCTTCCACTTACAGATGACCCCTCAATCCCATCATATATGATGCTC 1633  
Db 1586 GAACCAAAATTTCTTCCACTTACAGATGACCCCTCAATCCCATCATATATGATGCTC 1527  
QY 1634 CAGGAACCGGCATAGCCCGGTTTATTTGCTTCAACATTAAGAGAAATCCCAAGAAC 1693  
Db 1526 CAGGAACCGGCATAG--CCCGTTTATTTGCTTCAACATTAAGAGAAATCCCAAGAAC 1468  
QY 1694 AACAACCAATGAAATTTTGGAGCAATGTGTGTTTGTGCTGAGGATTAAGGATA 1753  
Db 1467 AACAACCAATGAAATTTTGGAGCAATGTGTGTTTGTGCTGAGGATTAAGGATA 1408  
QY 1754 GGGATTTATCTTACAGAAAGAGCTCAGACATTTCTTAAAGATGGATCTTAATCATC 1813  
Db 1407 GGGATTTATCTTACAGAAAGAGCTCAGACATTTCTTAAAGATGGATCTTAATCATC 1338  
QY 1814 TAAAGTTTCTCTCAAGAGATGCTCTGTGGGAGAGAGAAAGCCCAAGCAAGTATG 1873  
Db 1347 TAAAGTTTCTCTCAAGAGATGCTCTGTGGGAGAGAGAAAGCCCAAGCAAGTATG 1288  
QY 1874 TACAGACACATCCAGCTTCAATGCGCAGAGAGTGGGAAATTCCTCTCAGAGAGAAC 1933  
Db 1287 TACAGACACATCCAGCTTCAATGCGCAGAGAGTGGGAAATTCCTCTCAGAGAGAAC 1228  
QY 1934 GCCATATTTATGTGTGTGAGATGCAAGAAATATGCAAGATATGATATGATGCTCTT 1993  
Db 1227 GCCATATTTATGTGTGTGAGATGCAAGAAATATGCAAGATATGATATGATGCTCTT 1168

QY 1994 TCGAATATATAGCAAGAGGTTGAGTTGAAAACTAGAGCAATGAAAAACCTGCGCA 2053  
DB 1167 TCGAATATATAGCAAGAGGTTGAGTTGAAAACTAGAGCAATGAAAAACCTGCGCA 1108  
QY 2054 CTTTAAAGAGAAAAAGCTACCTTCAGGATTTTGGTCATA 2097  
DB 1107 CTTTAAAGAGAAAAAGCTACCTTCAGGATTTTGGTCATA 1064

RESULT 8  
BV178010/c 2933 bp DNA linear STS 10-JUN-2004  
LOCUS sqm97996 Human DNA (Sequenom) Homo sapiens STS genomic, sequence  
DEFINITION tagged site.  
ACCESSION BV178010 GI:48014252  
VERSION BV178010.1  
KEYWORDS STS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2933)  
Nelson,R.M., Marnellloe,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,  
Cantor,C.R. and Braun,A.  
Large-Scale Validation of Single Nucleotide Polymorphisms in Gene  
Regions  
Genome Res. (2004) In press

JOURNAL  
COMMENT Contact: Andreas Braun  
Pharmaceuticals division  
Sequenom, Inc.  
3595 John Hopkins Court, San Diego, CA 92121, USA  
Tel: 18582029018  
Fax: 18582029020  
Email: abraun@sequenom.com  
Primer A: No primer sequence submitted  
Primer B: No primer sequence submitted  
STS size: 2933.

FEATURES  
source Location/Qualifiers  
1..2933  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone\_lib="Human DNA (Sequenom)"  
<1..>2933

STS  
ORIGIN  
Query Match 63.8%; Score 1338; DB 11; Length 2933;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1778; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

DB 2546 GGTTTTGAACAAATGACAGTAAGCAACCAATCCATGTTGTAATTGAAACTTTGA 2487  
QY 675 GTCTCCTACCTACCGGTGGGTACCCCACTCAAGAGCTCTGTAATATTCCTGGTTT 734  
DB 2486 GTCTCCTACCTACCGGTGGGTACCCCACTCAAGAGCTCTGTAATATTCCTGGTTT 2427  
QY 735 ACCCCAGAAATATTTACAGATCACTGACAGAGTCTTGGCCAGAGGAAAGCCAAAT 794  
DB 2426 ACCCCAGAAATATTTACAGATCACTGACAGAGTCTTGGCCAGAGGAAAGCCAAAT 2367  
QY 795 ATCTGTACTTCAGCAGATCCAGTTTTCAGTGGCAATTTCAAGGCGATTCACTTAC 854  
DB 2366 ATCTGTACTTCAGCAGATCCAGTTTTCAGTGGCAATTTCAAGGCGATTCACTTAC 2307  
QY 855 TAGAATGATGCAATTAACCACTGCTGTGATGAAATTTGCAATTTCAATACAGACTT 914  
DB 2306 TAGAATGATGCAATTAACCACTGCTGTGATGAAATTTGCAATTTCAATACAGACTT 2247  
QY 915 TTCTCTACGCTGAGATGCTTACGCTGATCTGCCCTTAACAGTATTTGAGGTACA 974  
DB 2246 TTCTCTACGCTGAGATGCTTACGCTGATCTGCCCTTAACAGTATTTGAGGTACA 2187  
QY 975 AAGCTACTTCGAAGCTGACGCTTGAAGATTAAGAGAGCACTGCTCTTTGAAAT 1034  
DB 2186 AAGCTACTTCGAAGCTGACGCTTGAAGATTAAGAGAGCACTGCTCTTTGAAAT 2127  
QY 1035 AAGGCAAGACAAAGAAAGAGAGCTACTTACCCGCAATATCTGCGGAGATGTT 1094  
DB 2126 AAGGCAAGACAAAGAAAGAGAGCTACTTACCCGCAATATCTGCGGAGATGTT 2067  
QY 1095 TCTCCAGTTCATTTTACCTGAGTCTTGAATTCGAGCAATTCCTAAAGGCAATTTT 1154  
DB 2066 TCTCCAGTTCATTTTACCTGAGTCTTGAATTCGAGCAATTCCTAAAGGCAATTTT 2007  
QY 1155 GCGAGGCTTTGAGACTATACAGTACAGTCTGTAAGAGCGAGCTACAGAGCTGTG 1214  
DB 2006 GCGAGGCTTTGAGACTATACAGTACAGTCTGTAAGAGCGAGCTACAGAGCTGTG 1947  
QY 1215 CAGTAAACAAAGGAGCGAGCTATATAGCGCTTTGTAAGAGTCCGTGCTGTT 1274  
DB 1946 CAGTAAACAAAGGAGCGAGCTATATAGCGCTTTGTAAGAGTCCGTGCTGTT 1887  
QY 1275 GGAATCCTCTCTGCTTCCCTTCTTCCAGGCAACCACTCACTCTCTCTGAAATCT 1334  
DB 1886 GGAATCCTCTCTGCTTCCCTTCTTCCAGGCAACCACTCACTCTCTCTGAAATCT 1827  
QY 1335 TCTTAAACTTCAACCCAGACCATATTCGTGACAGCTCAAGTTATTTCAACCAGAAA 1394  
DB 1826 TCTTAAACTTCAACCCAGACCATATTCGTGACAGCTCAAGTTATTTCAACCAGAAA 1767  
QY 1395 GCTCCATTTTGTCTTCAATTTGTGAATTTCTGTACTGCAACAAGAGTTCTGG 1454  
DB 1766 GCTCCATTTTGTCTTCAATTTGTGAATTTCTGTACTGCAACAAGAGTTCTGG 1707  
QY 1455 GAAAGGATATGTAACAGCTGAGCTGAGCTTGTGATGCTTCAAGTTCTTCAAGCAAT 1514  
DB 1706 GAAAGGATATGTAACAGCTGAGCTGAGCTTGTGATGCTTCAAGTTCTTCAAGCAAT 1647  
QY 1515 ACATGATCCATGAAGACAGC-GGAAAGCCCTGAGCTCTAAGATATCATCTCTC 1573  
DB 1646 ACATGATCCATGAAGACAGC-GGAAAGCCCTGAGCTCTAAGATATCATCTCTC 1587  
QY 1574 GAAACCAAAATTTTCCACTTACAGAGAACCCCTCAATCCCAATCAATATGATGTC 1633  
DB 1586 GAAACCAAAATTTTCCACTTACAGAGAACCCCTCAATCCCAATCAATATGATGTC 1527  
QY 1634 CAGGAACCGGATAGCCCGTTATTTGGTTCTCAACATAGAGAAATCTCAAGAAC 1693  
DB 1526 CAGGAACCGGATAGC-CGCTTTATTTGGTTCTCAACATAGAGAAATCTCAAGAAC 1468  
QY 1694 AACACCCAGATGAATTTTGAAGCAATGATGTTTGTGCTGAGGCAATTAAGATTA 1753  
DB 1467 AACACCCAGATGAATTTTGAAGCAATGATGTTTGTGCTGAGGCAATTAAGATTA 1408

```
QY 1754 GGGATTATCTATTGAGAAAAGCTCAGACATTTCTTAAAGCATGGGATCTTAATCATC 1813
DB 1407 GGGATTATCTATTGAGAAAAGCTCAGATTTCTTAAAGCATGGGATCTTAATCATC 1348
QY 1814 TAAAGTTTCTCTTCAAGAGATGCTCTGTTGGGAGAGAGAAACCCCGAGAAAGTATG 1873
DB 1347 TAAAGTTTCTCTTCAAGAGATGCTCTGTTGGGAGAGAGAAACCCCGAGAAAGTATG 1288
QY 1874 TACAAGCAATCAGCTCATGCGCAGAGGAGGAGAGATCTCTCCAGAGAGACG 1933
DB 1287 TACAAGCAATCAGCTCATGCGCAGAGGAGGAGAGATCTCTCCAGAGAGACG 1228
QY 1934 GCCATATTTATGTTGTGAGATGCAAAAGATATGCGCAAGATGTAATGATGCTTGG 1993
DB 1227 GCCATATTTATGTTGTGAGATGCAAAAGATATGCGCAAGATGTAATGATGCTTGG 1168
QY 1994 TGCATAATTAAGAAAGAGGTTGAGATTTAAAAAATGAAAGCAATGAAACCTGGCCA 2053
DB 1167 TGCATAATTAAGAAAGAGGTTGAGATTTAAAAAATGAAAGCAATGAAACCTGGCCA 1108
QY 2054 CTTTAAAGAGAAAGAAAGCTACCTTCAGATATTTGGTCAATA 2097
DB 1107 CTTTAAAGAGAAAGAAAGCTACCTTCAGATATTTGGTCAATA 1064

RESULT 9
LOCUS BD077780 390 bp DNA linear PAT 27-AUG-2002
DEFINITION 5'EST of secretory protein in brain.
ACCESSION BD077780
VERSION BD077780.1 GI:22623383
KEYWORDS JP 2001512015-A/65.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 390)
AUTHORS Edwards,J.B.D.M., Duclet,A. and Lacroix,B.
TITLE 5'EST of secretory protein in brain
JOURNAL Patent: JP 2001512015-A 65 21-AUG-2001;
GENSET
COMMENT OS Homo sapiens (human)
PN JP 2001512015-A/65
PD 21-AUG-2001
PF 31-JUL-1998 JP 2000505293
PR 01-AUG-1997 US 08/905223
PI JEAN BAPTISTE DUMAS MILNE EDWARDS,ALYMERIC DUCLETT, BRUNO PI
LACROIX
PC C12N15/09,C07K14/47,C12M1/00,C12P21/02,C12Q1/68,C12N15/00 CC
Von Heljne matkix
CC score 6.9
CC seq SLISLASHSVSC/SN
FH key location/Qualifiers
FT sig_peptide 289..357.
Location/Qualifiers
FEATURES
source
1..390
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 18.4%; Score 386; DB 6; Length 390;
Best Local Similarity 100.0%; Pred. No. 4.5e-199;
Matches 386; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 970 GTACAAAGCTTACTCCAAAGACTGCAAGTTGAAGTAAAGAGAGCACTGGTCTTTTG 1029
DB 3 GTACAAAGCTTACTCCAAAGACTGCAAGTTGAAGTAAAGAGAGCACTGGTCTTTTG 62

QY 1030 AAAATTAAGGAGACACAAAGAAAGAGAGTACTTAAACCCAGCATATACCTGGGGA 1089
DB 63 AAAATTAAGGAGACACAAAGAAAGAGAGTACTTAAACCCAGCATATACCTGGGGA 122
```

```
QY 1090 TGTCTCTCAGTTCAATTTTAACTGTGTCTTGAATTCGAGCAATTCCTTAAAGGCA 1149
DB 123 TGTCTCTCAGTTCAATTTTAACTGTGTCTTGAATTCGAGCAATTCCTTAAAGGCA 182
QY 1150 TTTTTCGAGCCCTTGTGAGCTATACAGTGAACAGTCTGAAAAGGAGGAGGCTACAGAG 1209
DB 183 TTTTTCGAGCCCTTGTGAGCTATACAGTGAACAGTCTGAAAAGGAGGCTACAGAG 242
QY 1210 CTGTGAGTAAACAAGGAGGAGGAGGAGGAGTATAGCCGCTTTGTAGAGATGCTGTGCTGC 1269
DB 243 CTGTGAGTAAACAAGGAGGAGGAGGAGGAGTATAGCCGCTTTGTAGAGATGCTGTGCTGC 302
QY 1270 TTGTGAGTCTCTCTCTGCTTCCCTTCTTGGCAGCCACACTGATCTCTGCTCGAA 1329
DB 303 TTGTGAGTCTCTCTCTGCTTCCCTTCTTGGCAGCCACACTGATCTCTGCTCGAA 362
QY 1330 CATCTCTTAACTTCAACCCAGACC 1355
DB 363 CATCTCTTAACTTCAACCCAGACC 388

RESULT 10
LOCUS P121202S04 1353 bp DNA linear PRI 14-DEC-1999
DEFINITION Homo sapiens methionine synthase reductase (MTRR) gene, exon 5.
ACCESSION AF121205
VERSION AF121205.1 GI:6572530
KEYWORDS
SEGMENT 4 of 12
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1353)
AUTHORS Scherer,S.W. and Gravel,R.A.
TITLE Molecular cloning, expression and physical mapping of the human
JOURNAL methionine synthase reductase gene
MEDLINE Gene 240 (1), 75-88 (1999)
PUBMED 10564814
REFERENCE 2 (bases 1 to 1353)
AUTHORS Lecleerc,D.
TITLE Direct Submission
JOURNAL Submitted (20-JAN-1999) Human Genetics, Montreal Children's
Hospital, 4060 Ste-Catherine West, Montreal, Quebec H3Z 2Z3, Canada
FEATURES
source
1..1353
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/map="5p15.2-p15.3; between W11755 and D5S1957"
358..736
/feature="MTRR"
/number=5

ORIGIN
Query Match 18.2%; Score 381; DB 9; Length 1353;
Best Local Similarity 100.0%; Pred. No. 2.2e-196;
Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 401 GTTTAGAACTTGTGTGAGCCGTGATTTCTGCACTTGGCCAGCCCTCAGAAAGCAT 460
DB 357 GTTTAGAACTTGTGTGAGCCGTGATTTCTGCACTTGGCCAGCCCTCAGAAAGCAT 416
QY 461 TTAGGTCAAGCAGACAGACAAAGAGATTAAGTGGCGCATCTCCGGTGGCATCTTCAT 520
DB 417 TTAGGTCAAGCAGACAGACAAAGAGATTAAGTGGCGCATCTCCGGTGGCATCTTCAT 476
QY 521 CTTGAGGAGAGACTTGTGAAGTCAAGCTGTACATTTGAATCAATGATCGAGCTTC 580
```

Db 477 CCTTGAGCAGACCTTGTGAAGTCAGAGCTGCTACACATTTGATCTCAAGTCGACCTTC 536  
QY 581 TGAAGTTCATGATTCAGAGAAAGAGATTCGAGTTTGAAGCAAAATGCAGTAACA 640  
Db 537 TGAAGTTCATGATTCAGAGAAAGAGATTCGAGTTTGAAGCAAAATGCAGTAACA 596  
QY 641 GGAACCAATCCAAATGTTGTAATTTGAAGACTTGAAGTCTCACTTACCCGTCGGTACCC 700  
Db 597 GGAACCAATCCAAATGTTGTAATTTGAAGACTTGAAGTCTCACTTACCCGTCGGTACCC 656  
QY 701 CACTCTCACAAGCTCTCTGAATATTCCTGTTTACCCCAAGATATTTACAGTAATC 760  
Db 657 CACTCTCACAAGCTCTCTGAATATTCCTGTTTACCCCAAGATATTTACAGTAATC 716  
QY 761 TGCAGAGTCTCTGGCCAGG 781  
Db 717 TGCAGAGTCTCTGGCCAGG 737

RESULT 11  
AC010346

LOCUS AC010346 109626 bp DNA linear PRI 10-NOV-2000  
DEFINITION Homo sapiens chromosome 5 clone CITB-H1\_2018B2, complete sequence.  
AC010346  
VERSION AC010346.6 GI:11136705  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
TITLE 1 (bases 1 to 109626)  
JOURNAL DOB Joint Genome Institute and Stanford Human Genome Center.  
REFERENCE TITLE Unpublished  
JOURNAL 2 (bases 1 to 109626)  
REFERENCE TITLE DOB Joint Genome Institute.  
JOURNAL Direct Submission  
TITLE Submitted (15-SEP-1999) Production Sequencing Facility, DOB Joint  
JOURNAL Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
3 (bases 1 to 109626)  
REFERENCE TITLE DOB Joint Genome Institute and Stanford Human Genome Center.  
JOURNAL Direct Submission  
TITLE Submitted (10-NOV-2000) DOB Joint Genome Institute, 2800 Mitchell  
JOURNAL Drive, Walnut Creek, CA 94598, USA  
On Nov 10, 2000 this sequence version replaced gi:9256196.  
COMMENT Draft Sequence Produced by DOB Joint Genome Institute  
WWW.JGI.DOE.GOV  
WWW.SHGC.STANFORD.EDU  
Finishing Completed at Stanford Human Genome Center  
Quality: Phrap Quality >=40 99.9% of Sequence;  
STS Content:  
WI-9255 G05749.

## FEATURES

source

Location/Qualifiers  
1..109626  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/clone="CITB-H1\_2018B2"

## ORIGIN

Query Match 15.7%; Score 330; DB 9; Length 109626;  
Best Local Similarity 99.7%; Pred. No. 1.3e-168; Indels 0; Gaps 0;  
Matches 380; Conservative 0; Mismatches 1;  
QY 401 GTTTAGAACTTGTGTTGAGCCGTGATTCCTGAGCTTGGCCAGCCCTCAGAAACATT 460  
Db 88571 GTTTAGAACTTGTGTTGAGCCGTGATTCCTGAGCTTGGCCAGCCCTCAGAAACATT 88630  
QY 461 TTAGGTCAAGCAGAGCAAGAGAGATTAAGTGGCGCACTCCGGTGGCATCACTGGCAT 520  
Db 88631 TTAGGTCAAGCAGAGCAAGAGAGATTAAGTGGCGCACTCCGGTGGCATCACTGGCAT 88690

QY 521 CCTTGAGCAGACCTTGTGAAGTCAGAGCTGCTACACATTTGATCTCAAGTCGACCTTC 580  
Db 88691 CCTTGAGCAGACCTTGTGAAGTCAGAGCTGCTACACATTTGATCTCAAGTCGACCTTC 88750  
QY 581 TGAAGTTCATGATTCAGAGAAAGAGATTCGAGTTTGAAGCAAAATGCAGTAACA 640  
Db 88751 TGAAGTTCATGATTCAGAGAAAGAGATTCGAGTTTGAAGCAAAATGCAGTAACA 88810  
QY 641 GGAACCAATCCAAATGTTGTAATTTGAAGACTTGAAGTCTCACTTACCCGTCGGTACCC 700  
Db 88811 GGAACCAATCCAAATGTTGTAATTTGAAGACTTGAAGTCTCACTTACCCGTCGGTACCC 88870  
QY 701 CACTCTCACAAGCTCTCTGAATATTCCTGTTTACCCCAAGATATTTACAGTAATC 760  
Db 88871 CACTCTCACAAGCTCTCTGAATATTCCTGTTTACCCCAAGATATTTACAGTAATC 88930  
QY 761 TGCAGAGTCTCTGGCCAGG 781  
Db 88931 TGCAGAGTCTCTGGCCAGG 88951

RESULT 12  
AC025174

LOCUS AC025174 110756 bp DNA linear PRI 28-MAR-2002  
DEFINITION Homo sapiens chromosome 5 clone CTD-2072124, complete sequence.  
AC025174  
VERSION AC025174.5 GI:19774456  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
TITLE 1 (bases 1 to 110756)  
JOURNAL DOB Joint Genome Institute and Stanford Human Genome Center.  
REFERENCE TITLE Unpublished  
JOURNAL 2 (bases 1 to 110756)  
REFERENCE TITLE DOB Joint Genome Institute.  
JOURNAL Direct Submission  
TITLE Submitted (07-MAR-2000) Production Sequencing Facility, DOB Joint  
JOURNAL Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
3 (bases 1 to 110756)  
REFERENCE TITLE DOB Joint Genome Institute.  
JOURNAL Direct Submission  
TITLE Submitted (07-MAR-2002) Production Sequencing Facility, DOB Joint  
JOURNAL Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
4 (bases 1 to 110756)  
REFERENCE TITLE DOB Joint Genome Institute and Stanford Human Genome Center.  
JOURNAL Direct Submission  
TITLE Submitted (28-MAR-2002) DOB Joint Genome Institute, 2800 Mitchell  
JOURNAL Drive, Walnut Creek, CA 94598, USA  
On Mar 28, 2002 this sequence version replaced gi:19224767.  
COMMENT Draft Sequence Produced by DOB Joint Genome Institute  
WWW.JGI.DOE.GOV  
WWW.SHGC.STANFORD.EDU  
Finishing Completed at Stanford Human Genome Center  
Quality: Phrap Quality >=40 100% of Sequence;  
Estimated Total Number of Errors is 0.

## FEATURES

source

Location/Qualifiers  
1..110756  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/clone="CTD-2072124"

## ORIGIN

Query Match 15.7%; Score 330; DB 9; Length 110756;  
Best Local Similarity 99.7%; Pred. No. 1.3e-168; Indels 0; Gaps 0;  
Matches 380; Conservative 0; Mismatches 1;  
QY 401 GTTTAGAACTTGTGTTGAGCCGTGATTCCTGAGCTTGGCCAGCCCTCAGAAACATT 460

Db 20040 GTTTAGAAGCTTGGTTGAGCCGTGAGTTCCTGAGCTTCGGCCAGCCCTCAGAAAAGCATT 20099  
Qy 461 TTAGTCAAGCAGAGGAGCAAGAGAGAGATAGTGGCGACTCCGGTGGCATCACTGCAT 520  
Db 20100 TTAGTCAAGCAGAGGAGCAAGAGAGATAGTGGCGACTCCGGTGGCATCACTGCAT 20159  
Qy 521 CTTTGAAGCAGAGCCTTGTGAAGTCAAGGCTGCTACATTTGAATCTCAAGTGCAGCTTC 580  
Db 20160 CCTCGAGAGAGAGCCTTGTGAAGTCAAGGCTGCTACATTTGAATCTCAAGTGCAGCTTC 20219  
Qy 581 TGAATTCGAGTATTCAGGAAGAAAGATCTGAGTTTGAAGCAAAAGCAGTGACA 640  
Db 20220 TGAATTCGAGTATTCAGGAAGAAAGATCTGAGTTTGAAGCAAAAGCAGTGACA 20279  
Qy 641 GCAACCAATCCATGTGTATTTGAAGACTTGAAGCTTCACTACCCGTTCCGTAACCC 700  
Db 20280 GCAACCAATCCATGTGTATTTGAAGACTTGAAGCTTCACTACCCGTTCCGTAACCC 20339  
Qy 701 CACTTCACAGAGCTCTCTGATATTCCTGGTTTACCCCGAATATTTACAGGTACATC 760  
Db 20340 CACTTCACAGAGCTCTCTGATATTCCTGGTTTACCCCGAATATTTACAGGTACATC 20399  
Qy 761 TGCAGAGTCTCTTGGCCAGG 781  
Db 20400 TGCAGAGTCTCTTGGCCAGG 20420

RESULT 13  
AC022921 158199 bp DNA linear HTG 12-MAR-2000  
LOCUS AC022921  
DEFINITION Homo sapiens clone RP11-138P20, WORKING DRAFT SEQUENCE, 12  
unordered pieces.  
AC022921.2 GI:7229868  
VERSION AC022921.2  
KEYWORDS HTG; HTGS PHASE1; HTGS\_DRAFT.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 158199)  
Birtren,B., Linton,L., Nusbaum,C. and Lander,E.  
Homo sapiens, clone RP11-138P20  
Unpublished  
2 (bases 1 to 158199)  
Birtren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
Anderson,S., Baldwin,J., Barna,N., Beckert,R., Beda,F.,  
Boonstavily,L., Bouckgalter,B., Brown,A., Burkett,G., Castle,A.,  
Choepey,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,  
DeArrellano,K., Dewar,K., Domino,M., Doyle,M., Fennestor,J.,  
Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J.,  
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,  
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,  
Lander,E., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,  
Macdonald,P., Margulis,N., McEwan,P., McGurk,A., McKernan,K.,  
McPheters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,  
Norman,C.H., O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K.,  
Pierce,N., Pisanl,C., Pollara,V., Raymond,C., Riley,R., Rochman,D.,  
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,  
Stojanovic,N., Subramanian,A., Talamas,J., Teefay,S., Theodore,J.,  
Tirrell,A., Vasilev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,  
Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (07-FEB-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Mar 12, 2000 this sequence version replaced gi:6921909.  
All repeats were identified using RepeatMasker:  
Smit,A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information  
Center project name: L6314  
Center clone name: 138\_P 20  
----- Summary Statistics  
Sequencing vector: M13; M77815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 152636 bases at least Q40  
Consensus quality: 155474 bases at least Q30  
Consensus quality: 156388 bases at least Q20  
Insert size: 178000; agarose-fp  
Insert size: 157099; sum-of-contigs  
Quality coverage: 4.4 in Q20 bases; agarose-fp  
Quality coverage: 5.0 in Q20 bases; sum-of-contigs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 12 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\*  
1 1283: contig of 1283 bp in length  
\* 1284 1383: gap of 100 bp  
\* 1384 4203: contig of 2820 bp in length  
\* 4204 4303: gap of 100 bp  
\* 4304 6786: contig of 2483 bp in length  
\* 6787 6886: gap of 100 bp  
\* 6887 9683: contig of 2797 bp in length  
\* 9684 9783: gap of 100 bp  
\* 9784 12902: contig of 3119 bp in length  
\* 12903 13002: gap of 100 bp  
\* 13003 16429: contig of 3427 bp in length  
\* 16430 16529: gap of 100 bp  
\* 16530 25201: contig of 8672 bp in length  
\* 25202 25301: gap of 100 bp  
\* 25302 36759: contig of 11458 bp in length  
\* 36760 36859: gap of 100 bp  
\* 36860 53921: contig of 17062 bp in length  
\* 53922 54021: gap of 100 bp  
\* 54022 72055: contig of 18033 bp in length  
\* 72056 72154: gap of 100 bp  
\* 72155 102527: contig of 30373 bp in length  
\* 102528 102627: gap of 100 bp  
\* 102628 158199: contig of 55572 bp in length.  
location/Qualifiers  
1.158199  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="RP11-138P20"  
/clone\_1lb="RPC1-11 Human Male BAC"  
1.1283  
/note="assembly\_fragment"  
1384.4203  
/note="assembly\_fragment"  
4304.6786  
/note="assembly\_fragment"  
6887.9683  
/note="assembly\_fragment"  
clone\_end:SP6  
vector\_side:left  
9784.12902  
/note="assembly\_fragment"  
13003.16429  
/note="assembly\_fragment"  
16530.25201  
/note="assembly\_fragment"  
clone\_end:T7  
vector\_side:right  
25302.36759  
/note="assembly\_fragment"









Matches 379; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	401	GTTTAGAACTGNGGTGAGCCGGGATTCGTGACTCTGGCAGGCCCTCAGAAACATT	460
Db	83524	GTTTAAAGACTTGTGGTGTGAAGCCGTGATTCCTGACCTTGGCAGCCCTCAGAAACATT	83583
Qy	461	TTAGGTCAAGCAGAGGCAAGAAGAGTAATGTGGCGCATCCCGGTGGCATCCTTGAT	520
Db	83584	TTAGGTCAAGCAGAGGCAAGAAGAGTAATGTGGCGCATCCCGGTGGCATCCTTGAT	83643
Qy	521	CCTTGAGAGCAGACCTTGTGAAATCAGAGCTGCTACACATTTGAATCTCAAGTCGACCTTC	580
Db	83644	CCTTGAGAGCAGACCTTGTGAAATCAGAGCTGCTACACATTTGAATCTCAAGTCGACCTTC	83703
Qy	581	TGAGATTCGATGATTCAGAGAGAAAGATTCTGAGGTTTTGAAGCAAAATGCAGTGACA	640
Db	83704	TGAGATTCGATGATTCAGAGAGAAAGATTCTGAGGTTTTGAAGCAAAATGCAGTGACA	83763
Qy	641	GCAACCAATCCAAATGTTGTATTTGAAGACTTTGAGTCTTCACTTACCCGTTCCGTAACCC	700
Db	83764	GCAACCAATCCAAATGTTGTATTTGAAGACTTTGAGTCTTCACTTACCCGTTCCGTAACCC	83823
Qy	701	CACCTCTCAACAAGCTCTCTGAAATTTCTTGGTTTACCCCGAATATTTTACAGGTACATC	760
Db	83824	CACCTCTCAACAAGCTCTCTGAAATTTCTTGGTTTACCCCGAATATTTTACAGGTACATC	83883
Qy	761	TGCAGAGATCTCTTGCCAGG 781	
Db	83884	TGCAGAGATCTCTTGCCAGG 83904	

Search completed: August 27, 2005, 09:38:49  
Job time : 6013.21 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 23:18:31 / Search time 4546.04 Seconds  
(without alignments)  
17558.328 Million cell updates/sec

Title: US-09-371-347a-1

Perfect score: 2097  
Sequence: 1 atgagagaggttcgttact.....ttcagatattggtcacaata 2097

Scoring table: OLIGO\_NUC  
Gapop 60.0, Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size: 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

EST:  
1: gb\_est1:  
2: gb\_est2:  
3: gb\_est3:  
4: gb\_est4:  
5: gb\_est5:  
6: gb\_est6:  
7: gb\_est7:  
8: gb\_est8:  
9: gb\_est9:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1552	75.9	3100	3	BC062577 Homo sapi
2	956	45.6	3143	3	BC035977 Homo sapi
3	719	34.3	908	5	BC348674 BX348674
4	689	32.9	874	4	BM801462 AGENCOURT
5	623	29.7	646	7	CN260357 170004241
6	586	27.9	852	5	BO431497 AGENCOURT
7	565	26.9	565	1	AUT279788 AUT279788
8	543	25.9	877	1	AUT12440 AUT12440
9	531	25.3	1061	5	BQ218755 AGENCOURT
10	517	24.7	826	4	BI772430 603055786
11	512	24.4	776	6	CB997527 AGENCOURT
12	507	24.2	834	5	BU941078 AGENCOURT
13	470	22.4	521	6	CB164340 K-EST0225
14	455	21.7	822	1	AU132586
15	448	21.4	591	2	AM965709 EST37782
16	446	21.3	818	6	CD559384 AGENCOURT
17	434	20.7	591	4	BI025283 RCS-MT025
18	431	20.6	974	5	BX375211 BX375211
19	406	19.4	710	5	BU570323 AGENCOURT
20	384	18.3	527	4	BI025277 RCS-MT025
21	374	17.8	579	7	CN260360 170006001
22	367	17.5	642	2	BF346446 602020302
23	361	17.2	692	7	CN260359 170004706
24	360	17.2	685	4	BM049352 603626120

25	359	17.1	499	6	CD704108 EST20635
26	354	16.9	386	1	AA279726 z892410.r
27	351	16.7	839	4	BC531787 602560355
28	341	16.3	395	4	BM838530 K-EST0114
29	340	16.2	526	2	AM952883 EST364953
30	337	16.1	818	7	CF995233 AGENCOURT
31	335	16.0	413	2	BF810368 RCS-C1041
32	335	16.0	413	2	BF810479 RCS-C1014
33	332	15.8	366	1	AA085543 znd4h11.r
34	311	14.7	478	4	BM754488 K-EST0021
35	308	14.7	620	7	CK002453 AGENCOURT
36	302	14.4	440	4	BG877205 QV3-H1046
37	297	14.2	416	6	CB996520 AGENCOURT
38	292	13.9	528	2	BE301292 ba89B07.x
39	291	13.9	664	7	CR768694 DKE2D459K
40	291	13.9	667	7	CR770923 DKE2D469N
41	291	13.9	767	7	CR557482 DKE2D469K
42	276	13.2	481	7	CR549172 DKE2D459J
43	272	13.0	301	1	AL704780 DKE2D686M
44	264	12.6	366	6	CB298361 220019 re
45	257	12.3	366	2	BF808461 QV1-C1017

## ALIGNMENTS

RESULT 1	BC062577	3100 bp	mRNA	linear	HTC 25-NOV-2003
LOCUS	Homo sapiens	CDNA	clone IMAGE:5189058,	containing frame-shift errors.	
DEFINITION	BC062577.1	GI:38511756			
ACCESSION	BC062577				
VERSION	BC062577.1	GI:38511756			
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo;				
REFERENCE	1 (baaes 1 to 3100)				
AUTHORS	Strausberg, R.D., Collins, P.S., Wagner, L., Scheinen, C.M., Schlier, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhut, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, D., Hsieh, F., Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshnyuk, S., Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hilyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettelman, M., Madan, A.C., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, K., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywicki, M.I., Skalska, U., Small, D.E., Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.				
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
MEDLINE	22388257				
PUBMED	12477932				
REFERENCE	2 (baaes 1 to 3100)				
AUTHORS	Strausberg, R.				
TITLE	Direct Submision				
JOURNAL	Submitted (24-NOV-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>				
COMMENT	Contact: MGC help desk Email: <a href="mailto:cgabs-r@mail.nih.gov">cgabs-r@mail.nih.gov</a> Tissue Procurement: Life Technologies, Inc.				

CDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
Sequencing Center: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: nisc\_mgc@nigri.nih.gov  
Ahter,N., Ayale,K., Beckstrom-Sternberg,S.M., Benjamin,B.,  
Blakesley,R.M., Boufide,S., Breen,K., Brinkley,C., Brooks,S.,  
Dietsch,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,  
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legadi,R.,  
Maduro,Q.L., Mariello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,  
McDowell,J., Pearson,R., Stancirpop,S., Thomas,F.J., Touchman,J.W.,  
Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,  
Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>  
Series: IRAC Plate: 135 Row: e Column: 21  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA 91: 4505278  
This clone has the following problem: frame shifted.

## FEATURES

## source

1..3100  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAG:5189058"  
/tissue\_type="Colon, Kidney, Stomach, adult, whole pooled"  
/clone\_id="NIH\_MGC\_116"  
/lab\_host="DH10B"  
/note="Vector: PCMV-SPORT6"

## ORIGIN

Query Match 75.9%; Score 1592; DB 3; Length 3100;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1812; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
QY 283 GGTCTGGTATTCAGAAATACACTTCTTTCAGATGCGGGAAGATATGATTAACGA 342  
DB 172 GGTCTGGTATTCAGAAATACACTTCTTTCAGATGCGGGAAGATATGATTAACGA 231  
QY 343 CTTCAAGAGCTTGGAGCCCGGCATTTCTATGACACTGACATGACATGACTGTGAGT 402  
DB 232 CTTCAAGAGCTTGGAGCCCGGCATTTCTATGACACTGACATGACATGACTGTGAGT 291  
QY 403 TTGAAGCTTGTGAGCGGTGATTCGTGACACTGCGCAGCCCTCAGAAAGCATTTT 462  
DB 292 TTGAAGCTTGTGAGCGGTGATTCGTGACACTGCGCAGCCCTCAGAAAGCATTTT 351  
QY 463 AGGTCAAGCAGAGCAGAGAGATTAAGTGGCGCACTCCCGGTGCATCACTGCATCC 522  
DB 352 AGGTCAAGCAGAGCAGAGAGATTAAGTGGCGCACTCCCGGTGCATCACTGCATCC 411  
QY 523 TTGAAGCAGACCTTGTGAAGTCAAGCTGCTACATTTGAATTCAGATCGAGCTTCTG 582  
DB 412 TCAGAGCAGACCTTGTGAAGTCAAGCTGCTACATTTGAATTCAGATCGAGCTTCTG 471  
QY 583 AGATTGATTCAGAGAGAAAGATTCAGAGTTTGAAGCAAAATGAGTGAACAGC 642  
DB 472 AGATTGATTCAGAGAGAAAGATTCAGAGTTTGAAGCAAAATGAGTGAACAGC 531  
QY 643 AACCAATCCAAATGTTGTAATTAAGATTGAGTCTCACTTACCCGTCGGATCCCA 702  
DB 532 AACCAATCCAAATGTTGTAATTAAGATTGAGTCTCACTTACCCGTCGGATCCCA 591  
QY 703 CTCTCAAGAGCTTCTGGAATATTCCTGTTTACCCCAAGATTTTACAGATCATCTG 762  
DB 592 CTCTCAAGAGCTTCTGGAATATTCCTGTTTACCCCAAGATTTTACAGATCATCTG 651  
QY 763 CAGAGTCTCTTGGCAGAGAGAAAGCAAGTATCTGTGACTTCAGAGATTCAGTTT 822  
DB 652 CAGAGTCTCTTGGCAGAGAGAAAGCAAGTATCTGTGACTTCAGAGATTCAGTTT 711

QY 823 CAAGTCCAAATTTCAAAAGCAGTTCAACTTACTACGAATGATGCCATTAACCACTCTG 882  
DB 712 CAAGTCCAAATTTCAAAAGCAGTTCAACTTACTACGAATGATGCCATTAACCACTCTG 771  
QY 883 CTGTGTGAATTTGACATTTTCAAAATACAGACTTTTCTATCAGCTGGAATGCTTCAGC 942  
DB 772 CTGTGTGAATTTGACATTTTCAAAATACAGACTTTTCTATCAGCTGGAATGCTTCAGC 831  
QY 943 GTGATCGCCCTTAACAGTATTCAGAGTACAAAGCTTCTCAAGAGCTGACGTTGAA 1002  
DB 832 GTGATCGCCCTTAACAGTATTCAGAGTACAAAGCTTCTCAAGAGCTGACGTTGAA 891  
QY 1003 GATTAAGAGAGACCTGCGTCTTTTGAATTAAGGCAGACACAAGAAAGAAAGAGCT 1062  
DB 892 GATTAAGAGAGC-CTGCGTCTTTTGAATTAAGGCAGACACAAGAAAGAAAGAGCT 950  
QY 1063 ACCTTACCCAGATATACCTGCGGATGTTCTCTCAAGTCAATTTTACTGCTGCTT 1122  
DB 951 ACCTTACCCAGATATACCTGCGGATGTTCTCTCAAGTCAATTTTACTGCTGCTT 1010  
QY 1123 GAATCCGAGCAATTCCTAAAGAGATTTTTCGAGCCCTTGTGACTATACAGTAC 1182  
DB 1011 GAATCCGAGCAATTCCTAAAGAGATTTTTCGAGCCCTTGTGACTATACAGTAC 1070  
QY 1183 AGTCTGAAGAGCGCAGGCTACAGAGCTGTGAGTAAACAAGGGCAGCCGATTAAGC 1242  
DB 1071 AGTCTGAAGAGCGCAGGCTACAGAGCTGTGAGTAAACAAGGGCAGCCGATTAAGC 1130  
QY 1243 CGCTTTGTACAGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1302  
DB 1131 CGCTTTGTACAGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1190  
QY 1303 CAGCCACCACTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1362  
DB 1191 CAGCCACCACTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1250  
QY 1363 TGTGAGGCTCAAGTTTATTTACCCAGAGAAAGCTTCATTTGTCTTCAACATTTGAGAA 1422  
DB 1251 TGTGAGGCTCAAGTTTATTTACCCAGAGAAAGCTTCATTTGTCTTCAACATTTGAGAA 1310  
QY 1423 TTTCTGTCTACAGCCCAACAAGAGTTTCCGGAAGAGAGTATGTAACAGCTGCTGCGC 1482  
DB 1311 TTTCTGTCTACAGCCCAACAAGAGTTTCCGGAAGAGAGTATGTAACAGCTGCTGCGC 1370  
QY 1483 TTGTGATGCTTCAAGTTCTTTCAGCCCAACAATCATGATCCCATGAAGACAGCGGAAA 1542  
DB 1371 TTGTGATGCTTCAAGTTCTTTCAGCCCAACAATCATGATCCCATGAAGACAGCGGAAA 1430  
QY 1543 GCCCTGCTCTTAAGATATCATCTCTCTGAAACAACAATCTTTCCACTTACAGAT 1602  
DB 1431 GCCCTGCTCTTAAGATATCATCTCTCTGAAACAACAATCTTTCCACTTACAGAT 1490  
QY 1603 GACCCCTCAATCCCAATCAATGATGATGATGATGATGATGATGATGATGATGATGATG 1662  
DB 1491 GACCCCTCAATCCCAATCAATGATGATGATGATGATGATGATGATGATGATGATGATG 1550  
QY 1663 TTCTCAACAATGAGAGAACTCAAGAACCAACCCAGATGAAATTTTGGAGCAATG 1722  
DB 1551 TTCTCAACAATGAGAGAACTCAAGAACCAACCCAGATGAAATTTTGGAGCAATG 1610  
QY 1723 TGGTTGTTTTTGGCTGACAGCATTAAGATAGGAAATTAATTAATTAATTAATTAATTA 1782  
DB 1611 TGGTTGTTTTTGGCTGACAGCATTAAGATAGGAAATTAATTAATTAATTAATTAATTA 1670  
QY 1783 CATTTCTTAAGATGAGATCTTAATCATCTTAAGATGATGATGATGATGATGATGATGATG 1842  
DB 1671 CATTTCTTAAGATGAGATCTTAATCATCTTAAGATGATGATGATGATGATGATGATGATG 1730  
QY 1843 GTTGGAG 1902  
DB 1731 GTTGGAG 1790

[illegible]

Oy		6	GAAATGTGAGCAAGCTGGTGTACATAGAAATTTTCACAGATCTTGCCTGTATTAGACA	120
Db		112	GAAATATGTAGCAGACTGTGGTGTACATGAATTTTCACAGATCTTGCCTGTATTAGEMA	171
Oy		121	TCCGATAAGTATGACCTTAACACCGAAGCATCTCTGTGTGTGTGGTTCTACACG	180
Db		172	TCCGATAAGTATGACCTTAACACCGAAGCATCTCTGTGTGTGTGGTTCTACACG	231
Oy		181	GGCACCCGGAACCCACCCGACACAGCCCCGCAAGTTGTAAAGAAAATACGAACCAACA	240
Db		232	GGCACCCGGAACCCACCCGACACAGCCCCGCAAGTTGTAAAGAAAATACGAACCAACA	291
Oy		241	CTGCCGGTGTGATTTCTTGTGCTCACTCGCGGTATGGGTTACTGGGTCTCGGTATTCAGA	300
Db		292	CTGCCGGTGTGATTTCTTGTGCTCACTCGCGGTATGGGTTACTGGGTCTCGGTATTCAGA	351
Oy		301	TACACCTACTTGTGCAATGGGGGGAGAGATTAATGATTAACGACTTCAAGACTTGAGACC	360
Db		352	TACACCTACTTGTGCAATGGGGGGAGAGATTAATGATTAACGACTTCAAGACTTGAGACC	411
Oy		361	CGGCATTTTCTATGACACTGGAACATGCAATGATCTGTGTAGTTTAGAACCTTGGSTTAG	420
Db		412	CGGCATTTTCTATGACACTGGAACATGCAATGATCTGTGTAGTTTAGAACCTTGGSTTAG	471
Oy		421	CCGTTGATTTGCTGACCTGTGGCCAGCCCTCGAAGAAAGATTTTAGTCAAGCAGACAGCAA	480
Db		472	CCGTTGATTTGCTGACCTGTGGCCAGCCCTCGAAGAAAGATTTTAGTCAAGCAGACAGCAA	531
Oy		481	GAGAGATTAAGTGGCGCACTCCCGGTGGCATCACTTGATCTTTGAGACACAGCTTGTG	540
Db		532	GAGAGATTAAGTGGCGCACTCCCGGTGGCATCACTTGATCTTTGAGACACAGCTTGTG	591
Oy		541	AAGTCAGAGCTGTCAACATTTGAATCTCAAGTGGAGCTTCTGAGATTCAGATTCAGGA	600
Db		592	AAGTCAGAGCTGTCAACATTTGAATCTCAAGTGGAGCTTCTGAGATTCAGATTCAGGA	651
Oy		601	AGAAAGATTTCTGAGTTTTGGAAGCAAAATGCAAGTGAACAGCAACCAATCCAATGTGTGA	660
Db		652	AGAAAGATTTCTGAGTTTTGGAAGCAAAATGCAAGTGAACAGCAACCAATCCAATGTGTGA	711
Oy		661	ATTGAAGACTTTGAGTCTCACTTACCCGTTCCGGTACCCCACTTCAACAAGCTCTCTG	720
Db		712	ATTGAAGACTTTGAGTCTCACTTACCCGTTCCGGTACCCCACTTCAACAAGCTCTCTG	771
Oy		721	AATATTTCTGGTTTACCCCAGAAATATTTACAGGTAACTGTGACAGAGTCTTTGGCCAG	780
Db		772	AATATTTCTGGTTTACCCCAGAAATATTTACAGGTAACTGTGACAGAGTCTTTGGCCAG	831
Oy		781	GAGAAAGCCCAAGATCTGTGACTTCAAGCAGATCAATTTTTCAAGTCCCAATTTCAAG	840
Db		832	GAGAAAGCCCAAGATCTGTGACTTCAAGCAGATCAATTTTTCAAGTCCCAATTTCAAG	891
Oy		841	GCAGTCAACTTACTACGATGATGATGCCATTAACCACTCTGCTGTGAAATTTGACATT	900
Db		892	GCAGTCAACTTACTACGATGATGATGCCATTAACCACTCTGCTGTGAAATTTGACATT	951
Oy		901	TCAAAATCAAGCTTTTCTATCAAGCTTGGAGATGCCCTTCAAGCGATCTGCCCTTAACGT	960
Db		952	TCAAAATCAAGCTTTTCTATCAAGCTTGGAGATGCCCTTCAAGCGATCTGCCCTTAACGT	1011
Oy		961	GATTCTGAGGTACAAAGCTTACTCAAAAGACTGAGCTTGAAGATTAAGAAGACACTGCG	1020
Db		1012	GATTCTGAGGTACAAAGCTTACTCAAAAGACTGAGCTTGAAGATTAAGAAGACACTGCG	1071
Oy		1021	GTCCTTTGAAATTAAGGCGACACACAAAGAGAAGG	1058
Db		1072	GTCCTTTGAAATTAAGGCGACACACAAAGAGAAGG	1109

DEFINITION	EX348674 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
ACCESSION	EX348674
VERSION	EX348674.1
KEYWORDS	GI:30375301
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	unpublished (2001)
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primed end enriched, double-strand cDNA was digested with NotI and cloned into the NotI and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by life technologies, a division of Invitrogen. This sequence belongs to sequence cluster 3392.f For more information about this cluster, see http://www.genoscope.cns.fr/cdnat/e=CS0BAHG006ZB02_CS00490_1ec=3392.f
FEATURES	
source	Location/Qualifiers 1..908 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0DC010Y11" /cissue_type="NEUROBLASTOMA COT 25-NORMALIZED" /cissue_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED" /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primed end enriched, double-strand cDNA was digested with NotI and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN	
Query Match	34.3%; Score 719; DB 5; Length 908;
Best Local Similarity	100.0%; Pred. No. 0;
Matches	719; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	674 AGTCTCACTTACCCGTTGCGTACCCCACTCTCAAGCCTCTGTGAATATTCCTGGTT 733
Db	28 AGTCTCACTTACCCGTTGCGTACCCCACTCTCAAGCCTCTGTGAATATTCCTGGTT 87
QY	734 TACCCCCGAATATTTAAGATATCTGCAAGATCTCTTGGCCAGAGAAAGCCAAAG 793
Db	88 TACCCCCGAATATTTAAGATATCTGCAAGATCTCTTGGCCAGAGAAAGCCAAAG 147
QY	794 TATCTGTACTTCAGCAGATCCAGTTTTCAGTSCCAATTTCAAGGCAAGTCAACTTA 853
Db	148 TATCTGTACTTCAGCAGATCCAGTTTTCAGTSCCAATTTCAAGGCAAGTCAACTTA 207
QY	854 CTACGATGATGCTAAATAACCACTCTGCTGTGATGAAATTGGACATTTCAATACAGCT 913
Db	208 CTACGATGATGCTAAATAACCACTCTGCTGTGATGAAATTGGACATTTCAATACAGCT 267
QY	914 TTTCTTACAGCTGGAGATGCTTCAACCGTATCTGGCCCTTAACAGTATTTCTAGATAC 973
Db	268 TTTCTTACAGCTGGAGATGCTTCAACCGTATCTGGCCCTTAACAGTATTTCTAGATAC 327
QY	974 AAAGCTTACTCCAAAGACTGCAGCTGTGAAGATTAAGAGAGACAGTCGCTCTTTTGAAA 1033
Db	328 AAAGCTTACTCCAAAGACTGCAGCTGTGAAGATTAAGAGAGACAGTCGCTCTTTTGAAA 387
QY	1034 TAAAGGCAACAACAAGAAAGAGGCTTACCCCAAGATATTAAGCTGGGATGTT 1093
Db	388 TAAAGGCAACAACAAGAAAGAGGCTTACCCCAAGATATTAAGCTGGGATGTT 447
QY	1094 CTCCTCACTTATTTTACCTGTGTCTTGAATCCAGCAATTCCTTAAGAGCATTTT 1153

[illegible]

Db 110 GAAATATGTGACCAAGCTGTGTGATGATGATTTCTGACATCTTCACTGATTTAGTGA 169  
 Qy 121 TCCGATTAAGTATGACCTAAACCCGAAAGAGCTCCCTGTGTGTGTCTCTCAAG 180  
 Db 170 TCCGATTAAGTATGACCTAAACCCGAAAGAGCTCCCTGTGTGTGTGTCTCTCAAG 229  
 Qy 181 GGCACCGGAGACCCACCCGACACAGCCGCAAGTTGTGAAGAAATACAGAACCAACA 240  
 Db 230 GGCACCGGAGACCCACCCGACACAGCCGCAAGTTGTGAAGAAATACAGAACCAACA 289  
 Qy 241 CTGCGGATTTGATTTCTTCTGCTCACTGCGGTATGAGTTACTGAGTCTCGGTATTCAGA 300  
 Db 290 CTGCGGATTTGATTTCTTCTGCTCACTGCGGTATGAGTTACTGAGTCTCGGTATTCAGA 349  
 Qy 301 TACACTACTTTTGGCAATGGGGGGAAGATTAATTGAATTAACGATTCAGAGCTTGAGGC 360  
 Db 350 TACACTACTTTTGGCAATGGGGGGAAGATTAATTGAATTAACGATTCAGAGCTTGAGGC 409  
 Qy 361 CGGCAATTTCTATGACACTGACATGACATGATGATGATGATTTAGTGAACCTTGATTTGAG 420  
 Db 410 CGGCAATTTCTATGACACTGACATGACATGATGATGATTTAGTGAACCTTGATTTGAG 469  
 Qy 421 CGGTGATTTGCTGACCTGTGCGCAAGCCCTCAGAAAGCATTTTATGATCAAGACAGACAA 480  
 Db 470 CGGTGATTTGCTGACCTGTGCGCAAGCCCTCAGAAAGCATTTTATGATCAAGACAGACAA 529  
 Qy 481 GAGGAGATTAAGTGGGCACTCCCGTGGATCACTCGATCCTTGAGGACAGACCTTGTG 540  
 Db 530 GAGGAGATTAAGTGGGCACTCCCGTGGATCACTCGATCCTTGAGGACAGACCTTGTG 589  
 Qy 541 AAGTCAGAGCTCTACACATTTGATCTCAAGTCAGACTCTGAGATTCAGATTCAGAG 600  
 Db 590 AAGTCAGAGCTCTACACATTTGATCTCAAGTCAGACTCTGAGATTCAGATTCAGAG 649  
 Qy 601 AGAAAGATTTGAGGTTTGAAGCAAAATGCAAGTGAACAGACCAATCAATGTTGTA 660  
 Db 650 AGAAAGATTTGAGGTTTGAAGCAAAATGCAAGTGAACAGACCAATCAATGTTGTA 709  
 Qy 661 ATTGAAGATTTGAGTCTCACTTACCCGTTGGATACCCCACTCTCAGAGCTTCTG 720  
 Db 710 ATTGAAGATTTGAGTCTCACTTACCCGTTGGATACCCCACTCTCAGAGCTTCTG 769  
 Qy 721 AATATTCTGTGTTTACCC 740  
 Db 770 AATATTCTGTGTTTACCC 789

RESULT 5  
 CN260357 646 bp mRNA linear EST 16-MAY-2004  
 DEFINITION 17000424179730 GRN\_BS Homo sapiens cDNA 5', mRNA sequence.  
 ACCESSION CN260357  
 VERSION CN260357.1 GI:47276771  
 KEYWORDS EST.

SOURCE  
 ORGANISM  
 Homo sapiens (human)

REFERENCE  
 AUTHORS  
 Fukuyama, Y.; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 646)  
 Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,  
 Li, Y., Xu, C., Pang, R., Guegler, K., Rao, M.S., Mandalam, R.,  
 Lebkoweki, J. and Stanton, L.W.  
 Transcriptome characterization elucidates signaling networks that  
 control human ES cell growth and differentiation  
 Nat. Biotechnol. 22 (6), 707-716 (2004)  
 Contact: Brandenberger R  
 Regenerative Medicine  
 Geron Corporation  
 230 Constitution Drive, Menlo Park, CA 94025, USA  
 Tel: 650 473 8658  
 Fax: 650 473 7760  
 Email: rbrandenberger@geron.com  
 Insert Length: 646 Std Error: 0.00.

# FEATURES

source

Location/Qualifiers  
 1..646  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /tissue\_type="embryonic stem cells, cell lines H1, H7, and  
 H9"  
 /clone\_lib="GRN BS"  
 /note="Toigo AT primed, full-length enriched cDNA library  
 from undifferentiated hES cell lines H1 (p32), H7 (p29),  
 and H9 (p26) maintained in feeder-free conditions"

## ORIGIN

Query Match 29.7%; Score 623; DB 7; Length 646;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 623; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 987 AAGCTGACCTTGAAGATTAAGAGACAGCTGCTCTTTGAATAATTAAGCAGACAC 1046  
 Db 24 AAGCTGACCTTGAAGATTAAGAGACAGCTGCTCTTTGAATAATTAAGCAGACAC 83  
 Qy 1047 AAAGAAAGAGAGCTTACTTACCAGATTAATCTGCGGATTTCTCTGCAATTAAT 1106  
 Db 84 AAAGAAAGAGAGCTTACTTACCAGATTAATCTGCGGATTTCTCTGCAATTAAT 143  
 Qy 1107 TTTTACTGTGTCTTGAATCCGAGCAATTCCTAAAGAGATTTTTCGAGCCTTGT 1166  
 Db 144 TTTTACTGTGTCTTGAATCCGAGCAATTCCTAAAGAGATTTTTCGAGCCTTGT 203  
 Qy 1167 GGACTATACAGTGAAGTGTGAAAGCGAGGCTACAGAGCTGTGTGATTAACAG 1226  
 Db 204 GGACTATACAGTGAAGTGTGAAAGCGAGGCTACAGAGCTGTGTGATTAACAG 263  
 Qy 1227 GGCAGCCATTTATACCGCTTTGTACAGATGCTGTGCTCTTGGATTTCTCT 1286  
 Db 264 GGCAGCCATTTATACCGCTTTGTACAGATGCTGTGCTCTTGGATTTCTCTCT 323  
 Qy 1287 GCTTCCCTCTTTCGAGCCAGCACTCACTGCTCTGCTGCAACATCTTCTAACTTCA 1346  
 Db 324 GCTTCCCTCTTTCGAGCCAGCACTCACTGCTCTGCTGCAACATCTTCTTAACTTCA 383  
 Qy 1347 ACCAGACCATTAATGCTGTGCAAGTCAATTAATTTCAACAGAAAGCTCATTTGT 1406  
 Db 384 ACCAGACCATTAATGCTGTGCAAGTCAATTAATTTCAACAGAAAGCTCATTTGT 443  
 Qy 1407 CTTCACATTTGGAATTTCTGTCTACTGCAACACAGAGTTCTGCGAAGGAGTANG 1466  
 Db 444 CTTCACATTTGGAATTTCTGTCTACTGCAACACAGAGTTCTGCGAAGGAGTANG 503  
 Qy 1467 TACAGGCTGCGTGGCTGTTGCTTCACTGCTTCAAGCAATCATGATCCCA 1526  
 Db 504 TACAGGCTGCGTGGCTGTTGCTTCACTGCTTCAAGCAATCATGATCCCA 563  
 Qy 1527 TGAAGACAGCGGAAAGCCCTGCTCTTAAGATATCAATCTCTCTGCAACAAATTC 1586  
 Db 564 TGAAGACAGCGGAAAGCCCTGCTCTTAAGATATCAATCTCTCTGCAACAAATTC 623  
 Qy 1587 TTTCACTTACCAATGACCCCT 1609  
 Db 624 TTTCACTTACCAATGACCCCT 646

RESULT 6  
 BQ431497 852 bp mRNA linear EST 24-MAY-2002  
 LOCUS BQ431497  
 DEFINITION AGENCOURT 7894690 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:6158144  
 5', mRNA sequence.  
 ACCESSION BQ431497  
 VERSION BQ431497.1 GI:21170583  
 KEYWORDS EST.  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;





Db 241 AGTACTTACCCGACATATACCTCGGGATGTTCTTCAGTTTATCTGTG 300  
QY 1119 TCTTGAATCCGAGCAATTCCTTAAAGCATTTTGGCAGCCCTGTGACTATCCAG 1178  
Db 301 TCTTGAATCCGAGCAATTCCTTAAAGCATTTTGGCAGCCCTGTGACTATCCAG 360  
QY 1179 TGAAGTGTGAAAACGAGGCTACAGAGCTGTGAGTAAACAAGGGGAGCCGATTA 1238  
Db 361 TGAAGTGTGAAAACGAGGCTACAGAGCTGTGAGTAAACAAGGGGAGCCGATTA 420  
QY 1239 TACCGCTTTGTACGAGATGCTGTGCTGTGTGTGATCTCTCTGCTTCCCTTC 1298  
Db 421 TACCGCTTTGTACGAGATGCTGTGCTGTGTGTGATCTCTCTGCTTCCCTTC 480  
QY 1299 TTGCCAGCCACCACTAGTCTCTGCTGCAACATCTTCTTAACTTCAACCCAGACATA 1358  
Db 481 TTGCCAGCCACCACTAGTCTCTGCTGCAACATCTTCTTAACTTCAACCCAGACATA 540  
QY 1359 TTGCTGTGCAAGCTCAAGTTTATTT 1383  
Db 541 TTGCTGTGCAAGCTCAAGTTTATTT 565

RESULT 8  
AUI24440 877 bp mRNA linear EST 01-AUG-2002  
LOCUS AUI24440 NT2RM4 Homo sapiens cDNA clone NT2RM4000010 5', mRNA  
DEFINITION sequence.

ACCESSION AUI24440 GI:10949156  
VERSION AUI24440.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 877)  
Ota,T., Makamatsu,A., Ozawa,M., Ishii,S., Saito,K., Yamamoto,J.,  
Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S. and  
Ishigai,T.

TITLE HRI human cDNA project (Ota,T., Makamatsu,A., Ozawa,M., Ishii,S.,  
Saito,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T.,  
Suzuki,Y., Sugano,S., Isegai,T.)  
JOURNAL Unpublished (2000)  
COMMENT Contact: Takao Isegai  
Genomics Laboratory  
Helix Research Institute  
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986

FEATURES  
source  
1..877  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="NT2RM4000010"  
/cell\_type="teratocarcinoma"  
/cell\_line="NT2"  
/clone\_id="NT2RM4"  
/note="Vector: pME18SFL3; mRNA from uninduced NT2 neuronal  
precursor cells"

## ORIGIN

Query Match 25.9%; Score 543; DB 1; Length 877;  
Best Local Similarity 100.0%; Pred. No. 1.5e-286;  
Matches 543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1143 AAAGCATTTTTCGAGCCCTTGTGACATATACAGTGAAGGAGGAGGCT 1202  
|||||

Db 192 AAAGCATTTTTCGAGCCCTTGTGACATATACAGTGAAGGAGGAGGCT 251  
QY 1203 AAGAGAGCTGTGAGTAAACAAGGGGAGCCGATTAATACCCCTTGTAGAGATGCCG 1262  
Db 252 AAGAGAGCTGTGAGTAAACAAGGGGAGCCGATTAATACCCCTTGTAGAGATGCCG 311  
QY 1263 TGCCTGTGTTGGATATCTCTCTGCTTCCCTTCTTGGCCAGCCACCACTAGTCTCT 1322  
Db 312 TGCCTGTGTTGGATATCTCTCTGCTTCCCTTCTTGGCCAGCCACCACTAGTCTCT 371  
QY 1323 GCTGGAACATCTTCTTAACTTCAACCCAGACCATATGCTGTGCAAGCTCAAGTTAT 1382  
Db 372 GCTGGAACATCTTCTTAACTTCAACCCAGACCATATGCTGTGCAAGCTCAAGTTAT 431  
QY 1383 TCAACCAAGAAAGCTTCATTTTGTCTTCAACATTTGTGAATTTTGTCTACTGCCAGAC 1442  
Db 432 TCAACCAAGAAAGCTTCATTTTGTCTTCAACATTTGTGAATTTTGTCTACTGCCAGAC 491  
QY 1443 AGAGGTTCTGCGAAGGAGATATGTAAGGCTGTGCTGCTGCTTGTGTTGCTTCAAGTCT 1502  
Db 492 AGAGGTTCTGCGAAGGAGATATGTAAGGCTGTGCTGCTGCTTGTGTTGCTTCAAGTCT 551  
QY 1503 TCAGCCAAACATATGATGATCCCATGAAGACGCGAAAGCCCTGTGCTTCAAGTATC 1562  
Db 552 TCAGCCAAACATATGATGATCCCATGAAGACGCGAAAGCCCTGTGCTTCAAGTATC 611  
QY 1563 CATCTCTCTGCAACCAAAATTTCTTCACTTACCAATGACCCCTCAATCCCATCTCT 1622  
Db 612 CATCTCTCTGCAACCAAAATTTCTTCACTTACCAATGACCCCTCAATCCCATCTCT 671  
QY 1623 AATGTTGGTTCGAGAACCGGATAGCCCGTTTATGAGTTCTTCAACATAGAGAGAA 1682  
Db 672 AATGTTGGTTCGAGAACCGGATAGCCCGTTTATGAGTTCTTCAACATAGAGAGAA 731  
QY 1683 ACT 1685  
Db 732 ACT 734

RESULT 9  
BQ218755 1061 bp mRNA linear EST 02-MAY-2002  
LOCUS AGENCOURT 7565843 NIH\_MGC\_92 Homo sapiens cDNA clone IMAGE:6041670  
DEFINITION 5', mRNA sequence.  
ACCESSION BQ218755 GI:20400155  
VERSION BQ218755.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1061)  
NIH-MGC http://mhc.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
JOURNAL  
TITLE Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cgaabbs-remail.nih.gov  
Tissue Procurement: ATCC  
DNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1AM1379 row: n column: 07  
High quality sequence stop: 518.

## FEATURES

source  
1..1061  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6041670"  
/tissue\_type="embryonal carcinoma, cell line"

/lab host="DH10B (phage-resistant)"  
/clone lib="NIH\_MGC\_92"  
/note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 2.5 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 25.3%; Score 531; DB 5; Length 1061;  
Best Local Similarity 100.0%; Pred. No. 6,1e-280;  
Matches 531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 995 AGCTTGAAGATTAAGAGAGACACTGCGTCTTTGAAATTAAGAGACAAAGAGA 1054  
DB 1 AGCTTGAAGATTAAGAGAGACACTGCGTCTTTGAAATTAAGAGACAAAGAGA 60  
QY 1055 AAGAGACTACTTACCCACATTAACCTGCGGATTTCTTCCAGATTATTTTAACT 1114  
DB 61 AAGAGACTACTTACCCACATTAACCTGCGGATTTCTTCCAGATTATTTTAACT 120  
QY 1115 GGATCTTGAATTCGAGCAATTCCTAAAGGCAATTTTGCGAGCCCTTGTGACTATA 1174  
DB 121 GGATCTTGAATTCGAGCAATTCCTAAAGGCAATTTTGCGAGCCCTTGTGACTATA 180  
QY 1175 CCAAGTACAGTGTGAAAGCGAGGCTACAGAGCTGTGCAAGTAAACAAGGCGACCG 1234  
DB 181 CCAAGTACAGTGTGAAAGCGAGGCTACAGAGCTGTGCAAGTAAACAAGGCGACCG 240  
QY 1235 ATTTAATACCGCTTTTACAGATGCTGTGCTGTGCTTTGATCTCTCTGCTTCC 1294  
DB 241 ATTTAATACCGCTTTTACAGATGCTGTGCTGTGCTTTGATCTCTCTGCTTCC 300  
QY 1295 CTCTCTGCGACGACCACTAGTCTCTGCTGCAACATCTTCTTAACCTCAACCCAGAC 1354  
DB 301 CTCTCTGCGACGACCACTAGTCTCTGCTGCAACATCTTCTTAACCTCAACCCAGAC 360  
QY 1355 CATATTGCTGTGCAAGCTCAAGTTTATTTCAACCCAGAAAGCTTCATTTGTCTTACA 1414  
DB 361 CATATTGCTGTGCAAGCTCAAGTTTATTTCAACCCAGAAAGCTTCATTTGTCTTACA 420  
QY 1415 TTGTGGAATTTCTCTCTACAGTCCCAACAAGAGTTCTGCGAAGGAGTATGACAGCT 1474  
DB 421 TTGTGGAATTTCTCTCTACAGTCCCAACAAGAGTTCTGCGAAGGAGTATGACAGCT 480  
QY 1475 GGCTGCGCTTGTGTTGCTTCAAGTTCTTCAAGCCAAATATGATGCC 1525  
DB 481 GGCTGCGCTTGTGTTGCTTCAAGTTCTTCAAGCCAAATATGATGCC 531

RESULT 10 826 bp mRNA linear EST 25-SEP-2001  
BI772430 60305786F1 NIH\_MGC\_122 Homo sapiens cDNA clone IMAGE:5205285 5',  
DEFINITION mRNA sequence.  
ACCESSION BI772430.1 GI:15764008  
VERSION BI772430.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE NIH-MGC http://mgc.nci.nih.gov/  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Place: L1M1514 row: 1 column: 22  
High quality sequence stop: 824.  
Location/Qualifiers  
1..826  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5205285"  
/lab host="DH10B"  
/clone lib="NIH\_MGC\_122"  
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;  
Site 1: NotI; Site 2: EcoRV (destroyed); RNA source  
anonymous pool of 24 week female lung, 16 week female  
spleen, and 20-22 week male spleens. Library is oligo-dT  
primed and directionally cloned (EcoRV site is destroyed  
upon cloning). Average insert size 1.4 kb, insert size  
range 1-3 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 026. Note:  
this is a NIH\_MGC Library."

FEATURES

source

ORIGIN

Query Match 24.7%; Score 517; DB 4; Length 826;  
Best Local Similarity 99.7%; Pred. No. 3e-272;  
Matches 617; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ATGAGAGGTTCTGTACTATATGCTACACAGCAGGACAGCAAGGCCATGCGAGAA 60  
DB 53 ATGAGAGGTTCTGTACTATATGCTACACAGCAGGACAGCAAGGCCATGCGAGAA 112  
QY 61 GAAATGTGAGCAAGCTGTGTACATGATTTTCTGACATCTTCACTGATTAAGTAA 120  
DB 113 GAAATGTGAGCAAGCTGTGTACATGATTTTCTGACATCTTCACTGATTAAGTAA 172  
QY 121 TCCGATTAATGACTTAAACCCGAAACAGCTCTCTGTTGTGTTTCTACACG 180  
DB 173 TCCGATTAATGACTTAAACCCGAAACAGCTCTCTGTTGTGTTTCTACACG 232  
QY 181 GGCACGGAGACCCACCGCACAGCCCGCAAGTTGTTAAGAAATACAGAACCAACA 240  
DB 233 GGCACGGAGACCCACCGCACAGCCCGCAAGTTGTTAAGAAATACAGAACCAACA 292  
QY 241 CTGCCGATTATTTCTTTGCTCACTGCGGTATGAGTTACTGGGTCTCGGTATTCAGAA 300  
DB 293 CTGCCGATTATTTCTTTGCTCACTGCGGTATGAGTTACTGGGTCTCGGTATTCAGAA 352  
QY 301 TACACCTACTTTTGCATGCGGAGGAGATTAATGATTAACGACTTCAAGAGCTTGAAGCC 360  
DB 353 TACACCTACTTTTGCATGCGGAGGAGATTAATGATTAACGACTTCAAGAGCTTGAAGCC 412  
QY 361 CGGCATTTCATGACCTGGACATGACATGATGATCTGTAGGTTTAAACCTTGTGGTGA 420  
DB 413 CGGCATTTCATGACCTGGACATGACATGATGATCTGTAGGTTTAAACCTTGTGGTGA 472  
QY 421 CCGTGATTGCTGACTGTGGCCAGCCCTCAGAAACATTTTAAAGTCAAGAGACAA 480  
DB 473 CCGTGATTGCTGACTGTGGCCAGCCCTCAGAAACATTTTAAAGTCAAGAGACAA 532  
QY 481 GAGAGATTAAGTGGCGACCTCCGGTGGCATCACTGCACTTCTTGAGAGACACTTTGTG 540  
DB 533 GAGAGATTAAGTGGCGACCTCCGGTGGCATCACTGCACTTCTTGAGAGACACTTTGTG 592  
QY 541 AAGTCAGAGTGTACACATTAATCAAGTCAAGTCAAGTCTGAGATTTCATGATTCAGGA 600  
DB 593 AAGTCAGAGTGTACACATTAATCAAGTCAAGTCAAGTCTGAGATTTCATGATTCAGGA 652  
QY 601 AGAAGGATTCGAGGTTT 619  
DB 653 AGAAGGATTCGAGGTTT 671

RESULT 11  
 CB997527 776 bp mRNA linear EST 01-MAY-2003  
 LOCUS IMAGE:30338684 5', mRNA sequence.  
 DEFINITION CB997527.1 GI:30292047  
 ACCESSION CB997527  
 VERSION CB997527.1  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 776)  
 NIH-MGC http://mgi.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-r@mail.nih.gov  
 Tissue Procurement: Dr. Stefan Hannon  
 CDNA Library Preparation: Michael J. Brownstein (NHGRI) with help and advice from Piero Carninci (RIKEN)  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: NDAM365 row: 1 column: 21  
 High quality sequence stop: 564.  
 Location/Qualifiers  
 1..776  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30338684"  
 /tissue\_type="pre-ecclampic placenta"  
 /lab\_host="DH10B Tona"  
 /clone\_id="NIH MGC 148"  
 /note="Organ: placenta; Vector: pBluescriptR; Site 1: all-choi; Site 2: BamH; Library is oligo-dT primed and directionally cloned using primer 5'-TTTTTTTTTTTTTTN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH\_MGC Library."

# FEATURES source

## ORIGIN

Query Match 24.4%; Score 512; DB 6; Length 776;  
 Best Local Similarity 99.7%; Pred. No. 1.7e-269;  
 Matches 612; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGAGAGAGTTCTGTTACTATATGCTACAGCAGGAGGAGCAAGCCATCGAGAA 60  
 88 ATGAGAGAGTTCTGTTACTATATGCTACAGCAGGAGGAGCAAGCCATCGAGAA 147  
 61 GAATGTGTGAGCAAGCTGTGTACATGATTTTCTGCAGATCTTCACTGTAATTA 120  
 148 GAAATGTGTGAGCAAGCTGTGTACATGATTTTCTGCAGATCTTCACTGTAATTA 207  
 121 TCCGTAATATGACTTAATAAACCAGAAAGAGCTCTGTTGTTGTTCTTCAACAG 180  
 208 TCCGTAATATGACTTAATAAACCAGAAAGAGCTCTGTTGTTGTTCTTCAACAG 267  
 181 GGCACCGAGAGCCAGCCAGCAAGCCAGCAAGTTGTTAAGGAATATAGAACCAAA 240  
 268 GGCACCGAGAGCCAGCCAGCAAGCCAGCAAGTTGTTAAGGAATATAGAACCAAA 327  
 241 CTGCGGTTGATTTCTTTGCTCACTGCGGTATGGGTTACTGGGTTCTCGGTATTC 300  
 328 CTGCGGTTGATTTCTTTGCTCACTGCGGTATGGGTTACTGGGTTCTCGGTATTC 387

QY 301 TACACCTACTTTTGCAATGGGGGAGAGATATGATTAACAGCTTCAGAGCTTGAGCC 360  
 DB 388 TACACCTACTTTTGCAATGGGGGAGAGATATGATTAACAGCTTCAGAGCTTGAGCC 447  
 QY 361 CGGCAATTTTATGACACTGAGACATGCAATGACTGTGATGTTTGAAGTTGCTTGA 420  
 DB 448 CGGCAATTTTATGACACTGAGACATGCAATGACTGTGATGTTTGAAGTTGCTTGA 507  
 QY 421 CCGTGAATGCTGAGACTGCGCAGCCCTCAGAAAGCATTTTATGCTCAAGAGAGCA 480  
 DB 508 CCGTGAATGCTGAGACTGCGCAGCCCTCAGAAAGCATTTTATGCTCAAGAGAGCA 567  
 QY 481 GAGAGATATGAGGCGCACTCCCGGTGCACTGCATCTTGAAGACAGACCTTGTG 540  
 DB 568 GAGAGATATGAGGCGCACTCCCGGTGCACTGCATCTTGAAGACAGACCTTGTG 627  
 QY 541 AAGTCAAGCTGCTACACATTTCAATCTCAAGTCAAGCTTCTGAGATTCATTAAGA 600  
 DB 628 AAGTCAAGCTGCTACACATTTCAATCTCAAGTCAAGCTTCTGAGATTCATTAAGA 687  
 QY 601 AGAAGGATTTCTGA 614  
 DB 688 AGAAGGATTTCTGA 701

# FEATURES source

RESULT 12  
 BU941078 834 bp mRNA linear EST 18-OCT-2002  
 LOCUS IMAGE:6712893 5', mRNA sequence.  
 DEFINITION BU941078.1 GI:24129897  
 ACCESSION BU941078  
 VERSION BU941078.1  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 834)  
 NIH-MGC http://mgi.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-r@mail.nih.gov  
 Tissue Procurement: NCI  
 CDNA Library Preparation: Michael Brownstein Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LHC3022 row: e column: 21  
 High quality sequence stop: 586.  
 Location/Qualifiers  
 1..834  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6712893"  
 /tissue\_type="mixed (pool of 40 RNAs)"  
 /lab\_host="DH10B (Tl-phage-resistant)"  
 /clone\_id="NIH MGC 128"  
 /note="Vector: pDNR-LIB; Site 1: SfiI (ggcgcataagcc); Site 2: SfiI (ggcgcgcgcgcgc); Double-stranded cDNA was prepared from a pool of 40 cell line polyA+ RNAs (bladder - 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon - 4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%, kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were used in cloning as follows: 5'-AAGCAGTGTATCAAGCAGAGTGGCCTTACGCGCGG-3' and 5'-ATTCTAAGAGCCGAGCGCGCCCAATG-dT(30)NN-3'. Full-length enriched library was constructed using the Clontech

Creator SMART kit and size-selected to contain the >2 kb size fraction (other fractions present in NIH\_MGC\_126 and NIH\_MGC\_127). Library created in the laboratory of T. Uedini, M.D., Ph.D. (NIMH, NIH). Note: this is a NIH\_MGC library."

## ORIGIN

Query Match 24.2%; Score 507; DB 5; Length 834;  
Best Local Similarity 99.7%; Pred. No. 9,7e-267;  
Matches 607; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 379 GGCACATGCAATGATCTGTGTTAGTTGAACTTGTGTGAGCCGTGATGCTGCACTC 438  
Db 3 GGCATGCAATGATCTGTGTTAGTTGAACTTGTGTGAGCCGTGATGCTGCACTC 62

Qy 439 TGGCCAGCCCTCAGAAAGCAATTTAGTCAAGCAGAGACAGAGATTAAGTGGCGCA 498  
Db 63 TGGCCAGCCCTCAGAAAGCAATTTAGTCAAGCAGAGACAGAGATTAAGTGGCGCA 122

Qy 499 CTCGCCGTGATCAGCTGATCCTTGAGAGACAGACCTTGTGAAGTCAAGCTGCTAC 558  
Db 123 CTCGCCGTGATCAGCTGATCCTTGAGAGACAGACCTTGTGAAGTCAAGCTGCTAC 182

Qy 559 ATTGAATCTCAAGTCGAGCTTTGAGATTTGATGATTCAGAGAAAGATTTCTGAGTT 618  
Db 183 ATTGAATCTCAAGTCGAGCTTTGAGATTTGATGATTCAGAGAAAGATTTCTGAGTT 242

Qy 619 TTGAAGCAAAATGCAAGTGAAGCAAGCAATCCATGTTGTAATTGAAGCTTGAAGTCC 678  
Db 243 TTGAAGCAAAATGCAAGTGAAGCAAGCAATCCATGTTGTAATTGAAGCTTGAAGTCC 302

Qy 679 TCACTTACCCGTTGAGTACCCCACTCTCAAGCTCTGTGAATATCTGTGTTTACC 738  
Db 303 TCACTTACCCGTTGAGTACCCCACTCTCAAGCTCTGTGAATATCTGTGTTTACC 362

Qy 739 CCAAAATTTTACAGTATCATCTGCAAGATCTCTTGGCCAGAGAGAAACCAAGTCT 758  
Db 363 CCAAAATTTTACAGTATCATCTGCAAGATCTCTTGGCCAGAGAGAAACCAAGTCT 422

Qy 799 GTACCTCAGCAGATCCAGTTTTCAGAGTCAATTTCAAGGAGTTCAACTTACTACG 858  
Db 423 GTACCTCAGCAGATCCAGTTTTCAGAGTCAATTTCAAGGAGTTCAACTTACTACG 482

Qy 859 AATGATCCATTAATAAACCACTCTGCTGTGAATTTGACATTTCAATAACAGCTTTCC 918  
Db 483 AATGATCCATTAATAAACCACTCTGCTGTGAATTTGACATTTCAATAACAGCTTTCC 542

Qy 919 TATCAGCTGAGATGCTTCAAGGTGATCTGCCCTTAACAGTATTTGAGTGAACAAGC 978  
Db 543 TATCAGCTGAGATGCTTCAAGGTGATCTGCCCTTAACAGTATTTGAGTGAACAAGC 602

Qy 979 CTACTCCAA 987  
Db 603 CTACTCCAA 611

RESULT 13  
CB164340 521 bp mRNA linear EST 30-JAN-2003  
LOCUS K-E890225498 L17N670205n1 Homo sapiens cDNA clone  
DEFINITION L17N670205n1-39-F02 5', mRNA sequence.  
ACCESSION CB164340  
VERSION CB164340.1 GI:28150466  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 521)  
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.  
TITLE ZIC Frontier Korean EST Project 2001

## JOURNAL COMMENT

Unpublished (2002)  
Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 39 row: F column: 02  
High quality sequence stop: 521.

## FEATURES

## source

1. 521  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="L17N670205n1-39-F02"  
/sex="F"  
/lab\_host="top10F"  
/clone\_lib="L17N670205n1"  
/note="Organ: Liver; Vector: p77T3-Pac; Site 1: Scori; Site 2: NotI; The library was contributed by the Soares laboratory and it was constructed as described by Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome Research 6(9): 791-806. RNA was prepared from harvested cell culture."

## ORIGIN

Query Match 22.4%; Score 470; DB 6; Length 521;  
Best Local Similarity 99.8%; Pred. No. 2.1e-246;  
Matches 520; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1540 AAAGCCCTGCTCTTAAGATATCATCTCTCTGAAACAAATTTCTTCACTTACCA 1599  
Db 1 AAAGCCCTGCTCTTAAGATATCATCTCTCTGAAACAAATTTCTTCACTTACCA 60

Qy 1600 GATGACCCCTCATCCCCATCATTAATGATGATGATGATGATGATGATGATGATGAT 1659  
Db 61 GATGACCCCTCATCCCCATCATTAATGATGATGATGATGATGATGATGATGATGAT 120

Qy 1660 GGGTTCCTCAACATGAGAGAACTCAAGAACACACCCAGATGGAATTTTGGAGCA 1719  
Db 121 GGGTTCCTCAACATGAGAGAACTCAAGAACACACCCAGATGGAATTTTGGAGCA 180

Qy 1720 ATGTGTTGTTTGTGCTGACAGCATTAAGATGAGATTTATCTATTCAGAAAAAGCTC 1779  
Db 181 ATGTGTTGTTTGTGCTGACAGCATTAAGATGAGATTTATCTATTCAGAAAAAGCTC 240

Qy 1780 AGCATTTTCTTAAGCATGAGATCTTAATCATCTAAAGTTTCTTCTCAAGATGCT 1839  
Db 241 AGCATTTTCTTAAGCATGAGATCTTAATCATCTAAAGTTTCTTCTCAAGATGCT 300

Qy 1840 CTTGTTGGGAGGAGAGAGAGAGAGAGAGAGAGAGATATGATGATGATGATGATGATGAT 1899  
Db 301 CTTGTTGGGAGGAGAGAGAGAGAGAGAGAGAGAGATATGATGATGATGATGATGATGAT 360

Qy 1900 CAGCAGGTGCGAGAAATCTCTCTCAAGAGAAAGGCGCATATTTATGTTGAGATGCA 1959  
Db 361 CAGCAGGTGCGAGAAATCTCTCTCAAGAGAAAGGCGCATATTTATGTTGAGATGCA 420

Qy 1960 AAGAAATGCGCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2019  
Db 421 AAGAAATGCGCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480

Qy 2020 GTTGAATAACTAGAACATGAATAACCTTGCCACTTTTAA 2060  
Db 481 GTTGAATAACTAGAACATGAATAACCTTGCCACTTTTAA 521

RESULT 14  
AUI32586 822 bp mRNA linear EST 01-AUG-2002  
LOCUS AUI32586 NT2RPA Homo sapiens cDNA clone NT2RPA000141 5', mRNA  
DEFINITION sequence.

ACCESSION AU132586  
VERSION AU132586.1 GI:10929240  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 822)  
Ota, T., Sugiyama, T., Ishii, S., Suzuki, Y., Salto, K., Yamamoto, J.,  
Nishikawa, T., Nakamura, Y., Nagai, T., Sugano, S., Maehuo, Y. and  
Isegai, T.  
TITLE HRI human cDNA project (Ota, T., Sugiyama, T., Ishii, S., Suzuki, Y.,  
Salto, K., Yamamoto, J., Nishikawa, T., Nakamura, Y., Nagai, T.,  
Sugano, S., Maehuo, Y., Isegai, T.)  
JOURNAL Unpublished (2000)  
COMMENT Contact: Takao Isegai  
Genomics Laboratory  
Helix Research Institute  
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: genomics@hri.co.jp  
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix  
Research Institute; cDNA library construction: Department of  
Virology, Institute of Medical Science, University of Tokyo, and  
Helix Research Institute.  
FEATURES  
source 1..822  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="NT2RP4000141"  
/cell\_type="teratocarcinoma"  
/cell\_line="NT2"  
/clone\_1lb="NT2RP4"  
/note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor  
cells after 2-weeks retinoic acid (RA) induction"  
ORIGIN  
Query Match 21.7%; Score 455; DB 1; Length 822;  
Best Local Similarity 99.6%; Pred. No. 4.2e-238;  
Matches 555; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
283 GGTCTCGGATTCGAATACCTACTTTGCAATGGGGGAGATTAATGATTAAGCA 342  
182 GGTCTCGGATTCGAATACCTACTTTGCAATGGGGGAGATTAATGATTAAGCA 241  
343 CTTCAAGAGCTTGAGCCCGGCAATTTCTATGACATGACATGACATGATGATGATGAT 402  
242 CTTCAAGAGCTTGAGCCCGGCAATTTCTATGACATGACATGACATGATGATGATGAT 301  
403 TTAGAACTTGTGATGAGCCGCTGATGATGATGATGATGATGATGATGATGATGAT 462  
302 TTAGAACTTGTGATGAGCCGCTGATGATGATGATGATGATGATGATGATGATGAT 361  
463 AGTCAAGAGAGAGCAAGAGAGATTAAGTGGCGCATCCCGGTGGCATCACTGATCC 522  
362 AGTCAAGAGAGAGCAAGAGAGATTAAGTGGCGCATCCCGGTGGCATCACTGATCC 421  
523 TTGAGAGAGAGAGCTTGAGAGTCAAGAGCTGACATGATGATGATGATGATGATGAT 582  
422 TGGAGAGAGAGAGCTTGAGAGTCAAGAGCTGACATGATGATGATGATGATGATGAT 481  
583 AGATTGATGATTCAGAGAGAGAGATTTGAGAGTTTGAAGAGAGAGAGAGAGAGAG 642  
482 AGATTGATGATTCAGAGAGAGAGATTTGAGAGTTTGAAGAGAGAGAGAGAGAGAG 541  
643 AACCAATCAATGTTGATTAAGAGCTTGAGAGTCTCACTTACCCGTTGGATACCCCA 702  
542 AACCAATCAATGTTGATTAAGAGCTTGAGAGTCTCACTTACCCGTTGGATACCCCA 601  
703 CTCTCAAGAGCTCTCTGATATTTCTGTGTTTACCCCGAGATATTATTAAGGTATCTG 762

DB 602 CTCTCAAGAGCTCTCTGATATTTCTGTGTTTACCCCGAGATATTATTAAGGTATCTG 661  
QY 763 CAGAGATCTCTTGAG 822  
DB 662 CAGAGATCTCTTGAG 721  
QY 823 CAAGTCCCAATTTCAA 839  
DB 722 CAAGTCCCAATTTCAA 738  
RESULT 15  
AW965709  
LOCUS EST377782 MAGE resequences, MAGI Homo sapiens cDNA, mRNA sequence.  
ACCESSION AW965709  
KEYWORDS AW965709.1 GI:8155545  
SOURCE EST.  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 591)  
Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C.,  
Holt, I. E., Saeed, A. I., Sharov, V., Lee, N. H., Yeatman, T. J. and  
Quackenbush, J.  
TITLE Assessment of gene expression patterns in a model of colon tumor  
metastasis using a 19,200 element cDNA microarray  
JOURNAL Unpublished (2000)  
COMMENT Contact: John Quackenbush  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 3528  
Fax: 301 838 0208  
Email: johnqc@igr.org  
Plate: 218  
Seq primer: Reverse.  
FEATURES  
source 1..591  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_1lb="MAGE resequences, MAGI"  
/note="Vector: pBluescriptSKm"  
ORIGIN  
Query Match 21.4%; Score 448; DB 2; Length 591;  
Best Local Similarity 100.0%; Pred. No. 2.8e-234;  
Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1192 AAGCCAGGCTACAGAGAGTGTGACAGTAAACAAGGGGAGCGGATTAATGCGGCTTTGTA 1251  
DB 1 AAGCCAGGCTACAGAGAGTGTGACAGTAAACAAGGGGAGCGGATTAATGCGGCTTTGTA 60  
1252 CGAGATGCTGTGAGCTGTGATGATCTCCCTCGCTTCCCTTCCGACGACACA 1311  
DB 61 CGAGATGCTGTGAGCTGTGATGATCTCCCTCGCTTCCCTTCCGACGACACA 120  
1312 CTCAAGTCTCTGCTGGAACATCTTCTTAACCTTCAACCCAGACCAATATTGATGACAC 1371  
DB 121 CTCAAGTCTCTGCTGGAACATCTTCTTAACCTTCAACCCAGACCAATATTGATGACAC 180  
1372 TCAAGTTATTTTCAACCCAGAGAGCTCCATTTTGTCTTCAACATTTGGAATTTCTGTCT 1431  
DB 181 TCAAGTTATTTTCAACCCAGAGAGCTCCATTTTGTCTTCAACATTTGGAATTTCTGTCT 240  
1432 ACTGCAACAAGAGGTTCTGCGAAGGAGATGATGATGATGATGATGATGATGATGAT 1491  
DB 241 ACTGCAACAAGAGGTTCTGCGAAGGAGATGATGATGATGATGATGATGATGATGAT 300  
1492 GCTTCAAGTTCTTCAACCAATCAATGATGATGATGATGATGATGATGATGATGATGAT 1551  
DB 301 GCTTCAAGTTCTTCAACCAATCAATGATGATGATGATGATGATGATGATGATGATGAT 360

QY 1552 CCTAAGATATCGACTCTCTCCTGSAACAACAATTTTTCACCTTACCAAGATGACCCCTCA 1611  
Db 361 CCTAAGATATCGACTCTCTCCTGSAACAACAATTTTTCACCTTACCAAGATGACCCCTCA 420  
QY 1612 ATCCCATCATATGTGTGGGTCCAGGA 1639  
Db 421 ATCCCATCATATGTGTGGGTCCAGGA 448

Search completed: August 27, 2005, 15:58:31  
Job time : 4554.04 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 13:32:20 ; Search time 733.49 Seconds  
(without alignments)  
16924.161 Million cell updates/sec

Title: US-09-371-347A-1

Perfect score: 2097

Sequence: 1 atgaggaggttcctcgtact.....ttcagatatttgctacataa 2097

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*  
1: geneseqn19808:\*  
2: geneseqn19808:\*  
3: geneseqn20008:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004as:\*  
13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2097	100.0	3259	5	AA65070	AA65070 DNA encod
2	2097	100.0	3259	5	AA65070	AA65070 DNA encod
3	2097	100.0	3259	5	AA65070	AA65070 DNA encod
4	2097	100.0	3259	5	AA65070	AA65070 DNA encod
5	2097	100.0	3259	5	AA65070	AA65070 DNA encod
6	2097	100.0	3259	5	AA65070	AA65070 DNA encod
7	2097	100.0	3259	5	AA65070	AA65070 DNA encod
8	2097	100.0	3259	5	AA65070	AA65070 DNA encod
9	2097	100.0	3259	5	AA65070	AA65070 DNA encod
10	2097	100.0	3259	5	AA65070	AA65070 DNA encod
11	2097	100.0	3259	5	AA65070	AA65070 DNA encod
12	2097	100.0	3259	5	AA65070	AA65070 DNA encod
13	2097	100.0	3259	5	AA65070	AA65070 DNA encod
14	2097	100.0	3259	5	AA65070	AA65070 DNA encod
15	2097	100.0	3259	5	AA65070	AA65070 DNA encod
16	2097	100.0	3259	5	AA65070	AA65070 DNA encod
17	2097	100.0	3259	5	AA65070	AA65070 DNA encod
18	2097	100.0	3259	5	AA65070	AA65070 DNA encod
19	2097	100.0	3259	5	AA65070	AA65070 DNA encod
20	2097	100.0	3259	5	AA65070	AA65070 DNA encod

## ALIGNMENTS

21	279	13.3	591	12	ACH68540	ACH68540 Human gen
22	277	13.2	379	12	ACH82240	ACH82240 Human gen
23	225	10.7	503	5	AA65069	AA65069 DNA encod
24	188	9.0	525	12	ACH67438	ACH67438 Human gen
25	175	8.3	175	12	ACH81143	ACH81143 Human gen
26	158	7.5	2475	6	AA65071	AA65071 DNA encod
27	158	7.5	2475	13	AD161720	AD161720 Human gen
28	137	6.5	525	12	ACH73117	ACH73117 Human gen
29	124	5.9	175	12	ACH68648	ACH68648 Human gen
30	78	3.7	244	3	AA642736	AA642736 Human 5'
31	60	2.9	60	6	ABN36264	ABN36264 Human sp1
32	51	2.4	51	4	AA178548	AA178548 Human s11
33	38	1.8	1835	5	AA65071	AA65071 DNA encod
34	30	1.4	1681	11	AD131127	AD131127 Human CDN
35	26	1.2	26	3	AA58955	AA58955 PCR prime
36	26	1.2	26	3	AA58939	AA58939 PCR prime
37	26	1.2	26	6	ABX09549	ABX09549 Arteriosc
38	26	1.2	26	6	AA143713	AA143713 Pregestat
39	26	1.2	26	11	ADMA43205	ADMA43205 Human met
40	26	1.2	26	11	ADMA43189	ADMA43189 Human met
41	25	1.2	25	3	AA58952	AA58952 PCR prime
42	25	1.2	25	3	AA58937	AA58937 PCR prime
43	25	1.2	25	3	AA58947	AA58947 PCR prime
44	25	1.2	25	11	ADMA43187	ADMA43187 Human met
45	25	1.2	25	11	ADMA43202	ADMA43202 Human met

RESULT 1  
AA65070  
ID AA65070 standard; cDNA; 3259 BP.

AA65070;

13-FEB-2002 (first entry)

DNA encoding novel human diagnostic protein #874.

Human; chromosome mapping; gene mapping; gene therapy; forensic;

Food; supplement; medical imaging; diagnostic; genetic disorder; ss.

Homo sapiens.

WO2001/5067-A2.

30-MAR-2001; 2001WO-US008631.

31-MAR-2000; 2000US-00540217.

23-AUG-2000; 2000US-00649167.

(HYSE-) HYSEQ INC.

Dmanac RT, Liu C, Tang YT;

WPI; 2001-639362/73.

P-PSDB; ABG00883.

New isolated polynucleotide and encoded polypeptides, useful in

diagnostics, forensics, gene mapping, identification of mutations

responsible for genetic disorders or other traits and to assess

bioreactivity.

Claim 1; SEQ ID NO 874; 103bp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II)

sequences. (I) is useful as hybridisation probes, polymerase chain

reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

and in recombinant production of (II). The polynucleotides are also used

in diagnostics as expressed sequence tags for identifying expressed



CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantifying a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC diagnostics, forensics, gene mapping, identification of mutations in  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS4197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC fcp.wipo.int/pub/published\_pct\_sequences  
CC XX  
SQ Sequence 3259 BP; 944 A; 695 C; 674 G; 946 T; 0 U; 0 Other;  
Query Match 100.0%; Score 2097; DB 5; Length 3259;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAGAGGTTTCTGTTACTATATGCTACACAGAGGAGACAGGAAAGGCGCATGCGAGAA 60  
DB 80 ATGAGAGGTTTCTGTTACTATATGCTACACAGAGGAGACAGGAAAGGCGCATGCGAGAA 139  
QY 61 GAAATGTTGAGCAAGCTGTGTACATGATTTCTGAGATCTTCACTGATTTAGTAA 120  
DB 140 GAAATGTTGAGCAAGCTGTGTACATGATTTCTGAGATCTTCACTGATTTAGTAA 199  
QY 121 TCCGATTAAGTATGACTTAAACCGAAACAGCTCTCTTGTGTTGTGTTTCAACAGC 180  
DB 200 TCCGATTAAGTATGACTTAAACCGAAACAGCTCTCTTGTGTTGTGTTTCAACAGC 259  
QY 181 GGCACCGGAGACCAACCCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACA 240  
DB 260 GGCACCGGAGACCAACCCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACA 319  
QY 241 CTGCGGTTGATTTCTTGTCTGACCTGCGGTATGAGTTTACGAGTCTCGGATTTAGAA 300  
DB 320 CTGCGGTTGATTTCTTGTCTGACCTGCGGTATGAGTTTACGAGTCTCGGATTTAGAA 379  
QY 301 TACACCTACTTTTGAATGAGGAGGAGAAATTAATTAACGACTTCAAGAGCTTGAAGCC 360  
DB 380 TACACCTACTTTTGAATGAGGAGGAGAAATTAATTAACGACTTCAAGAGCTTGAAGCC 439  
QY 361 CGGATTTTATGACATCTGACATGACATGACATGATGATTTTGAAGCTTGTGAG 420  
DB 440 CGGATTTTATGACATCTGACATGACATGACATGATGATTTTGAAGCTTGTGAG 499  
QY 421 CCGTGAATTTGCTGACTCTGGCCAGCCCTCAGAAAGATTTTAAAGTCAAGAGAGACAA 480  
DB 500 CCGTGAATTTGCTGACTCTGGCCAGCCCTCAGAAAGATTTTAAAGTCAAGAGAGACAA 559  
QY 481 GAGAGATAAGTGGCGCACTCCGCGTGGCATCACCTGCATCTTGAAGACAGACCTTGTG 540  
DB 560 GAGAGATAAGTGGCGCACTCCGCGTGGCATCACCTGCATCTTGAAGACAGACCTTGTG 619  
QY 541 AAGTCAGAGCTGTACACATTTGAATTCATAGTGGAGTTTGAAGATTTGATTTAGGA 600  
DB 620 AAGTCAGAGCTGTACACATTTGAATTCATAGTGGAGTTTGAAGATTTGATTTAGGA 679  
QY 601 AGAAGAGATTCTGAGGTTTGAACAATAATGACAGAAACAGACCAATCCATGTTGTA 660  
DB 680 AGAAGAGATTCTGAGGTTTGAACAATAATGACAGAAACAGACCAATCCATGTTGTA 739  
QY 661 ATTGAAGACTTTGAGTCTCTCACTTAACCCGTTCCGATCCGCCACTCTCAAGACCTCTG 720  
DB 740 ATTGAAGACTTTGAGTCTCTCACTTAACCCGTTCCGATCCGCCACTCTCAAGACCTCTG 799  
QY 721 AATATCTCTGTTTACCCCGAATTTTACAGGTACATCTGAGAGTCTCTTGGCCAG 780

DB 800 AATATCTCTGTTTACCCCGAATAATTTACAGGTACATCTGAGAGTCTCTTGGCCAG 859  
QY 781 GAGAAAGCCAAATATCTGTGACTTGCAGAGATTCAGTTTTCATAGTGCATTTCAAG 840  
DB 860 GAGAAAGCCAAATATCTGTGACTTGCAGAGATTCAGTTTTCATAGTGCATTTCAAG 919  
QY 841 GCAATTCACTTACATGAGATGTCATTAACCACTCTGTGTGATTTGACATTT 900  
DB 920 GCAATTCACTTACATGAGATGTCATTAACCACTCTGTGTGATTTGACATTT 979  
QY 901 TCAATACAGACTTTTCCATACAGCTGAGAGATGCTTCAAGGTTATCTGCCCTAACG 960  
DB 980 TCAATACAGACTTTTCCATACAGCTGAGAGATGCTTCAAGGTTATCTGCCCTAACG 1039  
QY 961 GATTCTGAGTACAAAGCTTCTCAAGATGTCAGACTTGAAGATTAAGAGACACTGC 1020  
DB 1040 GATTCTGAGTACAAAGCTTCTCAAGATGTCAGACTTGAAGATTAAGAGACACTGC 1099  
QY 1021 GTCTTTTGAATAAAGGACACACAAAGAAAGAGCTTACCTTACCCAGCATTA 1080  
DB 1100 GTCTTTTGAATAAAGGACACACAAAGAAAGAGCTTACCTTACCCAGCATTA 1159  
QY 1081 CCGGCGGAGATGTTCTCTCAGATTCATTTTACCTGAGTCTTGAATCCGAGCATTCCT 1140  
DB 1160 CCGGCGGAGATGTTCTCTCAGATTCATTTTACCTGAGTCTTGAATCCGAGCATTCCT 1219  
QY 1141 AAAAAGCATTTTGTGCGAGCCCTTGTGACTATATACAGTACAGTCTGAAAAGCGCAG 1200  
DB 1220 AAAAAGCATTTTGTGCGAGCCCTTGTGACTATATACAGTACAGTCTGAAAAGCGCAG 1279  
QY 1201 CTACAGAGCTGTGACATTAACAGGGGCAAGCCGATTAATACCGCTTTGTACAGATGCC 1260  
DB 1280 CTACAGAGCTGTGACATTAACAGGGGCAAGCCGATTAATACCGCTTTGTACAGATGCC 1339  
QY 1261 TGTGCGGATGTTGTGATCTCTCTCCCTCGCTTCCCTTCCAGGACCAACATCAGTCTC 1320  
DB 1340 TGTGCGGATGTTGTGATCTCTCTCCCTCGCTTCCCTTCCAGGACCAACATCAGTCTC 1399  
QY 1321 CTGCTGAAACATCTTCTTAACTTCAACCCAGACATATTCGTGCAAGCTCAAGTTTA 1380  
DB 1400 CTGCTGAAACATCTTCTTAACTTCAACCCAGACATATTCGTGCAAGCTCAAGTTTA 1459  
QY 1381 TTTCACCCAGAAAGCTCCATTTTGTCTTCAACATTTGTGAATTTCTGTCTACTGCCACA 1440  
DB 1460 TTTCACCCAGAAAGCTCCATTTTGTCTTCAACATTTGTGAATTTCTGTCTACTGCCACA 1519  
QY 1441 ACAGAGTTCTGCGAAGGAGATGTAACAGCTGGCTGCTTGTGTGTTGCTTCACTT 1500  
DB 1520 ACAGAGTTCTGCGAAGGAGATGTAACAGCTGGCTGCTTGTGTGTTGCTTCACTT 1579  
QY 1501 CTTCAAGCAAACTATACATGATCCATGAAGACAGCGGGAAGCCCTGCTCTTAAGATA 1560  
DB 1580 CTTCAAGCAAACTATACATGATCCATGAAGACAGCGGGAAGCCCTGCTCTTAAGATA 1639  
QY 1561 TCCATCTCTCTGAAACACAAATTTCTTCACTTACCAATGACATCCCTCAATCCCATC 1620  
DB 1640 TCCATCTCTCTGAAACACAAATTTCTTCACTTACCAATGACATCCCTCAATCCCATC 1699  
QY 1621 ATATAGTGGGTCCAGAAACCGGCAATAGCCCGTTAATTTGGGTTCTTCAACAAATAGAG 1680  
DB 1700 ATATAGTGGGTCCAGAAACCGGCAATAGCCCGTTAATTTGGGTTCTTCAACAAATAGAG 1759  
QY 1681 AAATCCAGAACACACCCAGATGGAATTTTGAAGACATGTGTGTTTGTGCTGC 1740  
DB 1760 AAATCCAGAACACACCCAGATGGAATTTTGAAGACATGTGTGTTTGTGCTGC 1819  
QY 1741 AGGATTAAGATAGGATTAATCTATTCAGAAAAGAGCTCAGACATTTCTTAAAGCATGG 1800  
DB 1820 AGGATTAAGATAGGATTAATCTATTCAGAAAAGAGCTCAGACATTTCTTAAAGCATGG 1879  
QY 1801 ATCTTAATCATCTTAAGGTTTCTCTCAAGAGATGCTCTGTGGGAGAGAGAGCC 1860  
DB 1880 ATCTTAATCATCTTAAGGTTTCTCTCAAGAGATGCTCTGTGGGAGAGAGAGCC 1939



QY 1861 CCAGCAAGTATGTACAGACAACATCCAGCTTCAAGCCAGAGGCGAGAAATCCTC 1920  
DB 1940 CCAGAAAGTATGTACAGACAACATCCAGCTTCAAGCCAGAGGCGAGAAATCCTC 1999  
QY 1921 CTCAGAGAAAGCCCATATTTATGTGTGTGAGATGCAAGAAATATGCGCAAGATGA 1980  
DB 2000 CTCAGAGAAAGCCCATATTTATGTGTGTGAGATGCAAGAAATATGCGCAAGATGA 2059  
QY 1991 CATGATGCCCTTGTGCAATATATAGCAAGAGGTTGAGAGTTGAAAACTAGAACATG 2040  
DB 2060 CATGATGCCCTTGTGCAATATATAGCAAGAGGTTGAGAGTTGAAAACTAGAACATG 2119  
QY 2041 AAAACCCCTGGCATTAAAAAGAGAAAAAGCTACCTTCAGAGATTTTGGTCATTA 2097  
DB 2120 AAAACCCCTGGCATTAAAAAGAGAAAAAGCTACCTTCAGAGATTTTGGTCATTA 2176

RESULT 2  
AAC91226  
ID AAC91226 standard; DNA; 3259 BP.  
AC AAC91226;  
DT 20-MAR-2001 (first entry)  
XX Human schizophrenia related gene SEQ ID NO: 23.  
DE Human schizophrenia related gene SEQ ID NO: 23.  
XX Human; schizophrenia; developmental disorder; spina bifida cystica;  
KW Tourette's syndrome; bipolar illness; autism; conduct disorder;  
KW attention deficit disorder; obsessive compulsive disorder;  
KW chronic multiple tic syndrome; learning disorder; polymorphism; ds.  
XX Homo sapiens.  
OS Homo sapiens.  
PN WO200071754-A1.  
XX 30-NOV-2000.  
PD 24-MAY-2000; 2000WO-US014354.  
PF 25-MAY-1999; 99US-00318448.  
XX (UNNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.  
PA Johnson WG, Stenroos ES;  
PI WPI; 2001-025174/03.  
DR  
XX Diagnosing a developmental disorder, e.g. schizophrenia, by forming  
PT datasets (DS) of genetic (e.g. genotypes of folate metabolism alleles)  
PT and environmental variables affecting an individual and then comparing  
PT these DS with reference DS.  
XX  
PS Disclosure; Page 142-143; 156pp; English.  
CC The present invention provides a novel method of estimating the  
CC susceptibility of an individual to a developmental disorder using genetic  
CC and environmental variables. The method can be used in the diagnosis,  
CC prevention and treatment of disorders such as schizophrenia, spina bifida  
CC cystica, Tourette's syndrome, bipolar illness, autism, conduct disorders,  
CC attention deficit disorder, obsessive compulsive disorder, chronic  
CC multiple tic syndrome and learning disorders such as dyslexia  
CC  
SQ Sequence 3259 BP; 944 A; 695 C; 674 G; 946 T; 0 U; 0 Other;

Query Match 100.0%; Score 2097; DB 5; Length 3259;  
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Gaps 0;  
Matches 2097; Conservative 0; Indels 0; Gaps 0;

QY 1 ATGAGAGAGTTTCTGTATATATGTACAGAGAGGAGCAAGGCAATGCGAGAA 60  
DB 80 ATGAGAGAGTTTCTGTATATATGTACAGAGAGGAGCAAGGCAATGCGAGAA 139

QY 61 GAAATGTGTGAGAGCTGTGTACATGATTTTCTGAGATCTTCACTGTATTAGTAA 120  
DB 140 GAAATGTGTGAGAGCTGTGTACATGATTTTCTGAGATCTTCACTGTATTAGTAA 199  
QY 121 TCCGATTAAGTATGACCTTAAAAACGAAACAGCTCTCTGTGTGTGTGTTTCAACAG 180  
DB 200 TCCGATTAAGTATGACCTTAAAAACGAAACAGCTCTCTGTGTGTGTGTTTCAACAG 259  
QY 181 GGCACCGAGAGCCGACCCGAGACAGCCCGAGAGTTTGTAAAGAAATACAAACCAACA 240  
DB 260 GGCACCGAGAGCCGACCCGAGACAGCCCGAGAGTTTGTAAAGAAATACAAACCAACA 319  
QY 241 CTGCGCGGTGATTTCTTTGCTCCTGCGGTATGAGGTTA CTGGGTCTGCGGTATGAGG 300  
DB 320 CTGCGCGGTGATTTCTTTGCTCCTGCGGTATGAGGTTA CTGGGTCTGCGGTATGAGG 379  
QY 301 TACACCTAATTGTCATGAGGAGGAGAAAGTAAATTAACGACTTCAAGAGCTTGAACCC 360  
DB 380 TACACCTAATTGTCATGAGGAGGAGAAAGTAAATTAACGACTTCAAGAGCTTGAACCC 439  
QY 361 CGGCAATTTCTATGACATGAGACATGACAGATGACTGTGTAGGTTTGAACCTTGTGTAG 420  
DB 440 CGGCAATTTCTATGACATGAGACATGACAGATGACTGTGTAGGTTTGAACCTTGTGTAG 499  
QY 421 CCGTGGATTTGCTGACCTGCGCAGCCCTCAGAAAGCATTTTAAAGTCAAGAGAGAGACA 480  
DB 500 CCGTGGATTTGCTGACCTGCGCAGCCCTCAGAAAGCATTTTAAAGTCAAGAGAGAGACA 559  
QY 481 GAGAGATTAAGTGGGAGCATCCCGGTGAGCATCCTGATCCTTGAGAGACAGACTTGTG 540  
DB 560 GAGAGATTAAGTGGGAGCATCCCGGTGAGCATCCTGATCCTTGAGAGAGACTTGTG 619  
QY 541 AAGTCAGAGCTGCTACATGATTAATCTTAAGTCCAGCTTCTGAGATTGATGATTCAGGA 600  
DB 620 AAGTCAGAGCTGCTACATGATTAATCTTAAGTCCAGCTTCTGAGATTGATGATTCAGGA 679  
QY 601 AGAAAGATTTCTGAGGTTTGAAGCAAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660  
DB 680 AGAAAGATTTCTGAGGTTTGAAGCAAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 739  
QY 661 ATTGAAGACTTTGAGTCTCTCACTTACCCGTTGAGTCCGCACTCTCAAGAGCTCTGTG 720  
DB 740 ATTGAAGACTTTGAGTCTCTCACTTACCCGTTGAGTCCGCACTCTCTCAAGAGCTCTGTG 799  
QY 721 AATATTCCTGTTTACCCCGAGATTTTACAGTACATCTGACAGAGATCTCTTGGCCAG 780  
DB 800 AATATTCCTGTTTACCCCGAGATTTTACAGTACATCTGACAGAGATCTCTTGGCCAG 859  
QY 781 GAGGAAAGCCAGATCTGTGACTTACAGCAGATCCAGTTTTCAGAGGCAATTCAGAG 840  
DB 860 GAGGAAAGCCAGATCTGTGACTTACAGCAGATCCAGTTTTCAGAGGCAATTCAGAG 919  
QY 841 GCAGTTCAACTTACTACGATGATGATGCAATTAACCACTCTGCTGTGTAATGAGCAAT 900  
DB 920 GCAGTTCAACTTACTACGATGATGATGCAATTAACCACTCTGCTGTGTAATGAGCAAT 979  
QY 901 TCAATACAGACTTTTCTATACGCTGAGATGCTTACAGGATGATGCTTACAGAGT 960  
DB 980 TCAATACAGACTTTTCTATACGCTGAGATGCTTACAGGATGATGCTTACAGAGT 1039  
QY 961 GATTCTGAGTACAAAGCTTCTCAAGACTGACAGCTTGAAGATTAAGAGAGAGAGAGAG 1020  
DB 1040 GATTCTGAGTACAAAGCTTCTCAAGACTGACAGCTTGAAGATTAAGAGAGAGAGAGAG 1099  
QY 1021 GTCTCTTTGAAATTAAGGAGAGACAAAGAGAGAGAGAGTACCTTACCCAGCATTA 1080  
DB 1100 GTCTCTTTGAAATTAAGGAGAGACAAAGAGAGAGAGAGTACCTTACCCAGCATTA 1159  
QY 1081 CTTGCGGAGATGTTCTCTCAAGTTCAATTTTACCTGAGTCTTGAAGATCCAGAGCAATCTCT 1140  
DB 1160 CTTGCGGAGATGTTCTCTCAAGTTCAATTTTACCTGAGTCTTGAAGATCCAGAGCAATCTCT 1219

QY 1141 AAAAGGATTTTGGAGCCCTTGTGACTATACAGTGCAGTGTGTAAGGCGCAGG 1200  
 DB 1220 AAAAGGATTTTGGAGCCCTTGTGACTATACAGTGCAGTGTGTAAGGCGCAGG 1279  
 QY 1201 CTACAGAGCTGTGAGTAAACAAGGGGAGCGGATTTATAGCCGCTTTGAGAGATGCG 1260  
 DB 1280 CTACAGAGCTGTGAGTAAACAAGGGGAGCGGATTTATAGCCGCTTTGAGAGATGCG 1339  
 QY 1261 TGTGCTGTGTTGGATCTCTCTCTGCTTTTCTTGTGCGAGCAGCAGTCACTCTC 1320  
 DB 1340 TGTGCTGTGTTGGATCTCTCTCTGCTTTTCTTGTGCGAGCAGCAGTCACTCTC 1399  
 QY 1321 CTGTGGAACATCTTCTTAACTTCAACCCAGACATATTCGTGTGCAAGCTTAAGTTA 1380  
 DB 1400 CTGTGGAACATCTTCTTAACTTCAACCCAGACATATTCGTGTGCAAGCTTAAGTTA 1459  
 QY 1381 TTTCACCCAGGAAGGCTTCTTGTCTTCAATGTGGAATTTCTGTCTACTGCGACA 1440  
 DB 1460 TTTCACCCAGGAAGGCTTCTTGTCTTCAATGTGGAATTTCTGTCTACTGCGACA 1519  
 QY 1441 ACAGAGTTCTGCGAAGGAGATGTACAGGCTGCGCTTGTGTGTTGCTTCAATT 1500  
 DB 1520 ACAGAGTTCTGCGAAGGAGATGTACAGGCTGCGCTTGTGTGTTGCTTCAATT 1579  
 QY 1501 CTTCAGCCAAACATACATGATCCCATGAAAGACAGGGGAAAGCCCTGCTCTTAAGATA 1560  
 DB 1580 CTTCAGCCAAACATACATGATCCCATGAAAGACAGGGGAAAGCCCTGCTCTTAAGATA 1639  
 QY 1561 TCCATCTCTCTGGAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
 DB 1640 TCCATCTCTCTGGAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1699  
 QY 1621 ATATATGTTGGTTCAGGAACCGGCAATAGCCGTTTATTTGGTTTCTTCAACATAGAG 1680  
 DB 1700 ATATATGTTGGTTCAGGAACCGGCAATAGCCGTTTATTTGGTTTCTTCAACATAGAG 1759  
 QY 1681 AAATCTCAAGAAACAACCCAGATGGAATTTTGGAGCAATGTGTGTTTGGCTGC 1740  
 DB 1760 AAATCTCAAGAAACAACCCAGATGGAATTTTGGAGCAATGTGTGTTTGGCTGC 1819  
 QY 1741 AGGCTATAGGATAGGATTTATCTATTCAGAAAAGAGTCAAGATTTCTTAAAGCATGCG 1800  
 DB 1820 AGGCTATAGGATAGGATTTATCTATTCAGAAAAGAGTCAAGATTTCTTAAAGCATGCG 1879  
 QY 1801 ATCTTAATCATTAAGGTTTCTTCTCAAGATGCTCTGTGTGGGAGAGAGGCC 1860  
 DB 1880 ATCTTAATCATTAAGGTTTCTTCTCAAGATGCTCTGTGTGGGAGAGAGGCC 1939  
 QY 1861 CCAGCAAAATGTATCAAGCAACATCCAGCTTCAATGCGAGAGGTTGGAGAAATCTTC 1920  
 DB 1940 CCAGCAAAATGTATCAAGCAACATCCAGCTTCAATGCGAGAGGTTGGAGAAATCTTC 1999  
 QY 1921 CTCAGAGAAAGGCGCATTTATGTGTGTGAGATGCAAAAGATATGCGCAAGATGTA 1980  
 DB 2000 CTCAGAGAAAGGCGCATTTATGTGTGTGAGATGCAAAAGATATGCGCAAGATGTA 2059  
 QY 1981 CATGATGCTTGTGCAAAATATATAGCAAAAGAGTTGAGTTGAAAACTAGAAGCATG 2040  
 DB 2060 CATGATGCTTGTGCAAAATATATAGCAAAAGAGTTGAGTTGAAAACTAGAAGCATG 2119  
 QY 2041 AAAACCTGGGCACTTTAAAAAGAAAAACGCTACCTTCAAGATATTTGGTCATTA 2097  
 DB 2120 AAAACCTGGGCACTTTAAAAAGAAAAACGCTACCTTCAAGATATTTGGTCATTA 2176

RESULT 3  
 ID ADM43206 standard; cDNA; 3259 BP.  
 XX ADM43206;  
 XX AC  
 XX DT 03-JUN-2004 (first entry)  
 XX

DE Human full length cDNA encoding methionine synthase reductase.  
 XX  
 KW Human; ss; gene; Methionine synthase reductase polypeptide; HsmTRR;  
 KW cancer; cardiovascular disease; neural tube defect;  
 KW hyperhomocysteinemia; chromosome 5p15.2-p15.3; SNF;  
 KW single nucleotide polymorphism.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 80..2176  
 FT /tag= a "HsmTRR"  
 FT /product= "HsmTRR"  
 FT /replace(145,A)  
 FT /tag= b  
 FT /standard name= "Single\_nucleotide\_polymorphism"  
 FT /replace(189,A)  
 FT /tag= c  
 FT /standard\_name= "Single\_nucleotide\_polymorphism"  
 XX  
 PN US2003082676-A1.  
 XX  
 PD 01-MAY-2003.  
 XX  
 PF 10-AUG-1999; 99US-00371347.  
 XX  
 PR 16-JAN-1998; 98US-0071622P.  
 PR 15-JAN-1999; 99US-00232028.  
 XX  
 PA (GRAY/) GRAVEL R. A.  
 PA (ROZE/) ROZEN R.  
 PA (LECL/) LECLERC D.  
 PA (WILS/) WILSON A.  
 PA (ROSE/) ROSENBLATT D.  
 XX  
 PI Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;  
 XX  
 DR WPI; 2003-576610/54.  
 DR P-PSDB; ADM43207.  
 XX  
 PT New substantially pure nucleic acid encoding a mammalian methionine  
 PT synthase reductase polypeptide, useful for diagnosing, preventing or  
 PT treating conditions associated with altered methionine synthase activity,  
 PT e.g. cancer.  
 XX  
 PS Example 2; SEQ ID NO 24; 26pp; English.  
 PS  
 XX The invention relates to a substantially pure nucleic acid that encodes a  
 CC mammalian methionine synthase reductase polypeptide, HsmTRR, or that  
 CC hybridises at high stringency to a nucleic acid appearing as ADM43208 or  
 CC ADM43209. Also included are a non-human animal where one or both genetic  
 CC alleles encoding the methionine synthase reductase polypeptide are  
 CC mutated, an antibody that specifically binds the above methionine  
 CC synthase reductase polypeptide, a method of detecting the presence of the  
 CC above polypeptide, a method for detecting sequence variants for  
 CC methionine synthase reductase in a mammal, methods of treating or  
 CC preventing cancer (or cardiovascular disease or neural tube defects) in a  
 CC subject, methods of screening for a compound that modulates methionine  
 CC synthase reductase biological activity and a method for detecting an  
 CC increased risk of developing a neural tube defect in a mammalian embryo  
 CC or fetus. The nucleic acid is useful in diagnosing, preventing or  
 CC treating conditions associated with altered methionine synthase activity,  
 CC such as cancer, cardiovascular disease or neural tube defects, or in  
 CC screening for a compound that modulates methionine synthase reductase  
 CC biological activity. Naturally occurring variants of the polypeptide are  
 CC also associated with hyperhomocysteinemia. The gene for HsmTRR is  
 CC located on chromosome 5p15.2-p15.3. The present sequence is full length  
 CC sequence of the wild-type human hsmTRR cDNA.  
 CC  
 XX Sequence 3259 BP; 944 A; 695 C; 674 G; 946 T; 0 U; 0 Other;  
 SQ  
 Query Match 100.0%; Score 2097; DB 11; Length 3259;  
 Best Local Similarity 100.0%; Pred. No. 0;

Matches 2097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANAGAGAGTTTCTGTAATATGCTACACAGAGGACAGGCAAGCCATCGAGAA 60  
DB 80 ATGAGAGAGTTCTGTAATATGCTACACAGAGGACAGGCAAGCCATCGAGAA 139  
QY 61 GAAATGTGAGCACTGTGTACATGATTTTCTGAGATCTTCACTGATTAAGTAA 120  
DB 140 GAAATGTGAGCACTGTGTACATGATTTTCTGAGATCTTCACTGATTAAGTAA 199  
QY 121 TCCGATATGATGACCTTAAACCCGAAACAGCTCTCTTGTGTGTGTTCTACACG 180  
DB 200 TCCGATATGATGACCTTAAACCCGAAACAGCTCTCTTGTGTGTGTTCTACACG 259  
QY 181 GGCACCGGAGACCCACCCGACAGAGCCGCAAGTTTGTAAAGAAATPACGAACCAAC 240  
DB 260 GGCACCGGAGACCCACCCGACAGAGCCGCAAGTTTGTAAAGAAATPACGAACCAAC 319  
QY 241 CTGCGGTTGATTTCTTGTCTCACTGCGGTATGGGTTACTGGGTCTCGGTATTCAGAA 300  
DB 320 CTGCGGTTGATTTCTTGTCTCACTGCGGTATGGGTTACTGGGTCTCGGTATTCAGAA 379  
QY 301 TACACCTACTTTTGAATGAGGAGGAAATTAATGATTAACGACTTCAAGAGTTGAGCC 360  
DB 380 TACACCTACTTTTGAATGAGGAGGAAATTAATGATTAACGACTTCAAGAGTTGAGCC 439  
QY 361 CGGATTTCTATGACATGACATGACATGATGATGATGATTTTAAAGCTTGTGTGAG 420  
DB 440 CGGATTTCTATGACATGACATGACATGATGATGATGATTTTAAAGCTTGTGTGAG 499  
QY 421 CGGTGATTTCTGACATCTGCGGACGCTCAGAAAGATTTTAAAGTCAAGCAGAGGACA 480  
DB 500 CGGTGATTTCTGACATCTGCGGACGCTCAGAAAGATTTTAAAGTCAAGCAGAGGACA 559  
QY 481 GAGAGATTAAGTGGCGCATCTCCGATGATCATCTGATCTTGAAGACAGACTTGTG 540  
DB 560 GAGAGATTAAGTGGCGCATCTCCGATGATCATCTGATCTTGAAGACAGACTTGTG 619  
QY 541 AATGACAGCTGCTACATTTGAATCTCAAGTGTGATTTGAGATTTCCATGATTCAGGA 600  
DB 620 AATGACAGCTGCTACATTTGAATCTCAAGTGTGATTTGAGATTTCCATGATTCAGGA 679  
QY 601 AGAAGGATCTGAGGTTTGAAGCAAAATGACGTGAACAGCAACCAATCCAATGTTGTA 660  
DB 680 AGAAGGATCTGAGGTTTGAAGCAAAATGACGTGAACAGCAACCAATCCAATGTTGTA 739  
QY 661 ATTGAAGACTTTGATGCTCTCACTTACCCGTTGCTACCCCACTCTCAAGGCTCTG 720  
DB 740 ATTGAAGACTTTGATGCTCTCACTTACCCGTTGCTACCCCACTCTCAAGGCTCTG 799  
QY 721 AATATTCCTGTTTACCCCAAGATTTTACAGGTACATGCGAGAGTCTTTGGCCAG 780  
DB 800 AATATTCCTGTTTACCCCAAGATTTTACAGGTACATGCGAGAGTCTTTGGCCAG 859  
QY 781 GAGGAAAGCAAGATGTGTAAGCTTCAAGAGATCCAGTTTTCAGAGGCAATTTCAAG 840  
DB 860 GAGGAAAGCAAGATGTGTAAGCTTCAAGAGATCCAGTTTTCAGAGGCAATTTCAAG 919  
QY 841 GCAGTTCACTTACGAAATGATGCAATTAACCACTCTGCTGTGATTAAGCAAT 900  
DB 920 GCAGTTCACTTACGAAATGATGCAATTAACCACTCTGCTGTGATTAAGCAAT 979  
QY 901 TCAAAATCAAGCTTTTCTATCAAGCTGTGAGATGCTTCAAGGATCTGCTTAAAGT 960  
DB 980 TCAAAATCAAGCTTTTCTATCAAGCTGTGAGATGCTTCAAGGATCTGCTTAAAGT 1039  
QY 961 GATTTGAGGTAACAAAGCTATCTCAAAAGCTGAGATGCTTCAAGGATCTGCTTAAAGT 1020  
DB 1040 GATTTGAGGTAACAAAGCTATCTCAAAAGCTGAGATGCTTCAAGGATCTGCTTAAAGT 1099  
QY 1021 GTCTTTTGAATAAAGGACAGACAAAGAAAGAAAGCTTACCCAGCATATA 1080  
DB 1100 GTCTTTTGAATAAAGGACAGACAAAGAAAGAAAGCTTACCCAGCATATA 1159

QY 1081 CTGCGGAGATGTTCTCTCCAGTTCAATTTTACCTGGGTCTTGAATCCGAGCAATTCCT 1140  
DB 1160 CTGCGGAGATGTTCTCTCCAGTTCAATTTTACCTGGGTCTTGAATCCGAGCAATTCCT 1219  
QY 1141 AAAAAGCAATTTTGGAGCCCTTGTGACATATACAGTGAAGTGTGAAAAGCCAG 1200  
DB 1220 AAAAAGCAATTTTGGAGCCCTTGTGACATATACAGTGAAGTGTGAAAAGCCAG 1279  
QY 1201 CTACAGAGCTGTGACATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260  
DB 1280 CTACAGAGCTGTGACATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1339  
QY 1261 TGTGCTGTGTGATCTCTCCGCTTCCCTTCCCTTTCGACAGCACAACCTCACTCTC 1320  
DB 1340 TGTGCTGTGTGATCTCTCCGCTTCCCTTCCCTTTCGACAGCACAACCTCACTCTC 1399  
QY 1321 CTGCTGAAATCTTCTTAACTCAACCCAGACCATATGCTGTGCAAGCTCAAGTTTA 1380  
DB 1400 CTGCTGAAATCTTCTTAACTCAACCCAGACCATATGCTGTGCAAGCTCAAGTTTA 1459  
QY 1381 TTTCAACCGAAGAGTCAATTTTGTCTTCAATTTGAGAAATTTCTGTCTACTGCA 1440  
DB 1460 TTTCAACCGAAGAGTCAATTTTGTCTTCAATTTGAGAAATTTCTGTCTACTGCA 1519  
QY 1441 ACAGAGTTCTGCGAAGGAGATATGATACAGGCTGAGCTGCTGTGTGCTTCAAGTT 1500  
DB 1520 ACAGAGTTCTGCGAAGGAGATATGATACAGGCTGAGCTGCTGTGTGCTTCAAGTT 1579  
QY 1501 CTTCAGCCAAACATACATCATCCCATGAAGAAGGAGGAGGAGGAGGAGGAGGAGGAG 1560  
DB 1580 CTTCAGCCAAACATACATCATCCCATGAAGAAGGAGGAGGAGGAGGAGGAGGAGGAG 1639  
QY 1561 TCCATCTCTCTGAAACAAATTTTTCATCTTACAGATGACATCCCTCAATCCCATC 1620  
DB 1640 TCCATCTCTCTGAAACAAATTTTTCATCTTACAGATGACATCCCTCAATCCCATC 1699  
QY 1621 ATTAATGAGGATCAGAAACCCGATAGCCGCTTATTTGGGTTCTTCAACATAGAGAG 1680  
DB 1700 ATTAATGAGGATCAGAAACCCGATAGCCGCTTATTTGGGTTCTTCAACATAGAGAG 1759  
QY 1681 AAATCTCAAGAACCAACCCAGATGAGAAATTTTGAAGCAATGTGTTTGTGCTGC 1740  
DB 1760 AAATCTCAAGAACCAACCCAGATGAGAAATTTTGAAGCAATGTGTTTGTGCTGC 1819  
QY 1741 AGGATTAAGATAGGATTAATCTATTCAGAAAGAGCTCAAGATTTCTTAAAGATG 1800  
DB 1820 AGGATTAAGATAGGATTAATCTATTCAGAAAGAGCTCAAGATTTCTTAAAGATG 1879  
QY 1801 ATCTTAATCTATTAAGGTTTCTCTCAAGAGATGCTCTGTGAGGAGGAGGAGGAGG 1860  
DB 1880 ATCTTAATCTATTAAGGTTTCTCTCAAGAGATGCTCTGTGAGGAGGAGGAGGAGG 1939  
QY 1861 CCAGCAAGATATGACAGCAACATCAAGCTTCAAGGCTCAAGGAGTGGAGAAATCTCTC 1920  
DB 1940 CCAGCAAGATATGACAGCAACATCAAGCTTCAAGGCTCAAGGAGTGGAGAAATCTCTC 1999  
QY 1921 CTCAGAGAACGAGCAATTTATGTGTGTGAGATGCAAGAAATATGCGCAAGATGTA 1980  
DB 2000 CTCAGAGAACGAGCAATTTATGTGTGTGAGATGCAAGAAATATGCGCAAGATGTA 2059  
QY 1981 CATGATGCTGTGTGCAATTAATTAAGCAAGAGGTTGAGTGAAGAACTGAAAGCAATG 2040  
DB 2060 CATGATGCTGTGTGCAATTAATTAAGCAAGAGGTTGAGTGAAGAACTGAAAGCAATG 2119  
QY 2041 AAAAACCCTGCGCACTTTAAAGAAAGAAAGCGTCACTTCAAGATATTTGTGATATA 2097  
DB 2120 AAAAACCCTGCGCACTTTAAAGAAAGAAAGCGTCACTTCAAGATATTTGTGATATA 2176

RESULT 4  
ADM43208  
ID ADM43208 standard; cDNA; 2094 BP.

XX ADM43208;  
AC 03-JUN-2004 (first entry)  
DT  
XX  
XX Human wild-type methionine synthase reductase CDS.  
DE  
XX  
XX Human; 88; Methionine synthase reductase polypeptide; HsMTR; cancer;  
KW cardiovascular disease; neural tube defect; hyperhomocysteinaemia;  
KW chromosome 5p15.2-p15.3; SNP; single nucleotide polymorphism.  
XX  
XX Homo sapiens.  
OS  
FH Key Location/Qualifiers  
FT CDS 1..2094  
FT /tag= a  
FT /product= "hsmtrr"  
FT /partial  
FT /note= "No stop codon shown"  
FT replacement(66,A)  
FT /tag= b  
FT /standard\_name= "Single\_nucleotide\_polymorphism"  
FT /replace(110,A)  
FT /tag= c  
FT /standard\_name= "Single\_nucleotide\_polymorphism"  
XX US2003082676-A1.  
XX 01-MAY-2003.  
XX 10-AUG-1999; 99US-00371347.  
XX  
XX 16-JAN-1998; 98US-0071622P.  
XX 15-JAN-1999; 99US-00232028.  
XX  
XX (GRAV/) GRAVEL R A.  
XX (ROZE/) ROZEN R.  
XX (LECL/) LECLEERC D.  
XX (WILS/) WILSON A.  
XX (ROSE/) ROSENBLATT D.  
XX  
XX Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;  
XX WPI; 2003-576610/54.  
XX P-PSDB; ADM43207.  
XX  
XX New substantially pure nucleic acid encoding a mammalian methionine  
XX synthase reductase polypeptide, useful for diagnosing, preventing or  
XX treating conditions associated with altered methionine synthase activity,  
XX e.g. cancer.  
XX  
XX Claim 3; SEQ ID NO 1; 26pp; English.  
XX  
XX The invention relates to a substantially pure nucleic acid that encodes a  
XX mammalian methionine synthase reductase polypeptide, HsMTRR, or that  
XX hybridizes at high stringency to a nucleic acid appearing as ADM43208 or  
XX ADM43209. Also included are a non-human animal where one or both genetic  
XX alleles encoding the methionine synthase reductase polypeptide are  
XX mutated, an antibody that specifically binds the above methionine  
XX synthase reductase polypeptide, a method of detecting the presence of the  
XX above polypeptide, a method for detecting sequence variants for  
XX methionine synthase reductase in a mammal, methods of treating or  
XX preventing cancer (or cardiovascular disease or neural tube defects) in a  
XX subject, methods of screening for a compound that modulates methionine  
XX synthase reductase biological activity and a method for detecting an  
XX increased risk of developing a neural tube defect in a mammalian embryo  
XX or foetus. The nucleic acid is useful in diagnosing, preventing or  
XX treating conditions associated with altered methionine synthase activity,  
XX such as cancer, cardiovascular disease or neural tube defects, or in  
XX screening for a compound that modulates methionine synthase reductase  
XX biological activity. Naturally occurring variants of the polypeptide are  
XX also associated with hyperhomocysteinaemia. The gene for HsMTRR is  
XX located on chromosome 5p15.2-p15.3. The present sequence is the coding

CC sequence of the wild-type human hsmtrr cDNA.  
XX  
SQ Sequence 2094 BP; 591 A; 489 C; 481 G; 533 T; 0 U; 0 Other;  
Query Match 99.9%; Score 2094; DB 11; Length 2094;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2094; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAGAGGTTTCTGTACTATATGCTATACACAGAGGACAGCAAGCCATCGAGAA 60  
DB 1 ATGAGAGGTTTCTGTACTATATGCTATACACAGAGGACAGCAAGCCATCGAGAA 60  
QY 61 GAAATGTGAGCAACCTGTGTACATGATTTTCTGAGATCTTACGTATTAGTAA 120  
DB 61 GAAATGTGAGCAACCTGTGTACATGATTTTCTGAGATCTTACGTATTAGTAA 120  
QY 121 TCCGATATGATGACCTTAAACCCGAAAGAGCTCTCTTGTGTGTGTTTCTACAG 180  
DB 121 TCCGATATGATGACCTTAAACCCGAAAGAGCTCTCTTGTGTGTGTTTCTACAG 180  
QY 181 GGCACCGAGACCCACCCGACACAGCCCGCAAGTTGTTAAGAAATACAAACAA 240  
DB 181 GGCACCGAGACCCACCCGACACAGCCCGCAAGTTGTTAAGAAATACAAACAA 240  
QY 241 CTGCCGTTGATTTCTTGTCTCATCTGCGGTATAGGTTACTGGTCTCGGTATGAA 300  
DB 241 CTGCCGTTGATTTCTTGTCTCATCTGCGGTATAGGTTACTGGTCTCGGTATGAA 300  
QY 301 TACACCTACTTTTGCATGAGGGGAGATATATGATTAACGACTTCAGAGCTTGA 360  
DB 301 TACACCTACTTTTGCATGAGGGGAGATATATGATTAACGACTTCAGAGCTTGA 360  
QY 361 CGGCATTTCTATGACACTGACATGCAATGATGACTGTGATGTTTGAACCTTGT 420  
DB 361 CGGCATTTCTATGACACTGACATGCAATGATGACTGTGATGTTTGAACCTTGT 420  
QY 421 CCGTGATTTGCTGACCTGCGCCAGCCCTCAGAAACATTTTATGATCAAGAGG 480  
DB 421 CCGTGATTTGCTGACCTGCGCCAGCCCTCAGAAACATTTTATGATCAAGAGG 480  
QY 481 GAGAGATTAAGTGGCCACTCCCGTGGCATCACTGATCTTGTGAGGACAGACTT 540  
DB 481 GAGAGATTAAGTGGCCACTCCCGTGGCATCACTGATCTTGTGAGGACAGACTT 540  
QY 541 AAGTCAGAGCTGTACACATTTGATCTCAAGTCAAGTCTTGAATTCATGATTC 600  
DB 541 AAGTCAGAGCTGTACACATTTGATCTCAAGTCAAGTCTTGAATTCATGATTC 600  
QY 601 AGAAGGATTCGAGGTTTTTGAAGCAAAATGACGTGAACAGCAACCAATCAAT 660  
DB 601 AGAAGGATTCGAGGTTTTTGAAGCAAAATGACGTGAACAGCAACCAATCAAT 660  
QY 661 ATTGAAGACTTTGAGTCTCTCACTTACCCGTTCCGATCCCACTTCACAGCC 720  
DB 661 ATTGAAGACTTTGAGTCTCTCACTTACCCGTTCCGATCCCACTTCACAGCC 720  
QY 721 AATATTCCTGTTTACCCCAAGATATTTTACAGATCATCTGACAGAGTCTT 780  
DB 721 AATATTCCTGTTTACCCCAAGATATTTTACAGATCATCTGACAGAGTCTT 780  
QY 781 GAGAAAGCCAAATCTGTGACCTTCAAGAGATTCAGGTTTCAAGGCCAATTT 840  
DB 781 GAGAAAGCCAAATCTGTGACCTTCAAGAGATTCAGGTTTCAAGGCCAATTT 840  
QY 841 GCAGTTCAACTTCTACGATGATGCAATTAACCACTCTGCTGATGATTTGA 900  
DB 841 GCAGTTCAACTTCTACGATGATGCAATTAACCACTCTGCTGATGATTTGA 900  
QY 901 TCAAAATACAGACTTTTCTATACGCTTGAAGATGCTTCAAGCTATCTCC 960  
DB 901 TCAAAATACAGACTTTTCTATACGCTTGAAGATGCTTCAAGCTATCTCC 960  
QY 961 GATTTCAGATACAAAGCCATCTCAAAAGCTGACGTTGAATTAAGAGGAC 1020

Db 961 GATTCGAGGTACAAAGCCTACTCCAAAGCTGCAGCTTGAATATAAGAGAGCACTGC 1020  
 QY 1021 GTTCCTTTGAAAAATAAGGACACACAAAGAGAAAGAGCTACTTACCACGACATATA 1080  
 Db 1021 GTTCCTTTGAAAAATAAGGACACACAAAGAGAAAGAGCTACTTACCACGACATATA 1080  
 QY 1081 CCGGGGAGATGTTCTCCAGTTCATTTTACCTGGTGTGAAATCCGAGCAATTCCT 1140  
 Db 1081 CCGGGGAGATGTTCTCCAGTTCATTTTACCTGGTGTGAAATCCGAGCAATTCCT 1140  
 QY 1141 AAAAAGGCAATTTTGGAGCCCTTGTGACTATACAGTGCAGTCTGAAAAAGCCAG 1200  
 Db 1141 AAAAAGGCAATTTTGGAGCCCTTGTGACTATACAGTGCAGTCTGAAAAAGCCAG 1200  
 QY 1201 CTACAGAGCTGTGCAGTAAACAAGGGGAGCCGATTAATAGCCGTTTGTACAGATGCC 1260  
 Db 1201 CTACAGAGCTGTGCAGTAAACAAGGGGAGCCGATTAATAGCCGTTTGTACAGATGCC 1260  
 QY 1261 TGTGCTGCTGTGTGATCTCCCTCTGCTTCCCTTCCGAGGACACACTCAGTCTC 1320  
 Db 1261 TGTGCTGCTGTGTGATCTCCCTCTGCTTCCCTTCCGAGGACACACTCAGTCTC 1320  
 QY 1321 CTGCTGAAATCTTCTTAACTTCAACGACATATTCGTGTCAAGTCAAGTTTA 1380  
 Db 1321 CTGCTGAAATCTTCTTAACTTCAACGACATATTCGTGTCAAGTCAAGTTTA 1380  
 QY 1381 TTTCACCCAGAAAAGCTCCATTTTGTCTTCAACATTTGTGATGATTTCTGTCTACCCACA 1440  
 Db 1381 TTTCACCCAGAAAAGCTCCATTTTGTCTTCAACATTTGTGATGATTTCTGTCTACCCACA 1440  
 QY 1441 ACGAGGTTCTGGGAGGAGATATGACAGGCTGGCCCTTGTGTGTTGCTTCAATT 1500  
 Db 1441 ACGAGGTTCTGGGAGGAGATATGACAGGCTGGCCCTTGTGTGTTGCTTCAATT 1500  
 QY 1501 CTTGACCCAAACATACATGATCCCATGAGACAGGAGGAAACCTGTGCTCTTAAGATA 1560  
 Db 1501 CTTGACCCAAACATACATGATCCCATGAGACAGGAGGAAACCTGTGCTCTTAAGATA 1560  
 QY 1561 TCCATCTCTCTCGAACAACAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
 Db 1561 TCCATCTCTCTCGAACAACAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
 QY 1621 ATTAATGTTGGTTCAGAAACCGGACATAGCCGTTTATTTGGTTCTTACAACTAAGAG 1680  
 Db 1621 ATTAATGTTGGTTCAGAAACCGGACATAGCCGTTTATTTGGTTCTTACAACTAAGAG 1680  
 QY 1681 AAATCTCAAGAAACAACCCAGATGGAATTTTGAAGCAATGTTGTTTGTGCTGC 1740  
 Db 1681 AAATCTCAAGAAACAACCCAGATGGAATTTTGAAGCAATGTTGTTTGTGCTGC 1740  
 QY 1741 AGGATTAAGATAGGATTAATCTATTCAAAAAAGCTCAGACATTTCTTAAAGATGG 1800  
 Db 1741 AGGATTAAGATAGGATTAATCTATTCAAAAAAGCTCAGACATTTCTTAAAGATGG 1800  
 QY 1801 ATCTTAATCATTTAAAGTTTCTTCTCAAGAGATGCTCTCTTGGGAGAGAGAGCC 1860  
 Db 1801 ATCTTAATCATTTAAAGTTTCTTCTCAAGAGATGCTCTCTTGGGAGAGAGAGCC 1860  
 QY 1861 CCAGAAAGTATATACAAACATCAGCTTCATGAGCAGAGAGGAGGAGATCCTC 1920  
 Db 1861 CCAGAAAGTATATACAAACATCAGCTTCATGAGCAGAGAGGAGGAGATCCTC 1920  
 QY 1921 CTCAGAGAGAGCGGCATATTTATGTTGTGAGATGCAAAAGATATGCGCAAGATGA 1980  
 Db 1921 CTCAGAGAGAGCGGCATATTTATGTTGTGAGATGCAAAAGATATGCGCAAGATGA 1980  
 QY 1981 CAGGATGCCCTTTGCAAAATTAAGCAAGAGTGTGAATTAAGAACTAAGAGCAATG 2040  
 Db 1981 CAGGATGCCCTTTGCAAAATTAAGCAAGAGTGTGAATTAAGAACTAAGAGCAATG 2040  
 QY 2041 AAAACCTGGCCTTTAAAGAGAAAAGCTTACCTTCAAGATATTTGTGCA 2094  
 Db 2041 AAAACCTGGCCTTTAAAGAGAAAAGCTTACCTTCAAGATATTTGTGCA 2094

Db 2041 AAAACCTGGCCTTTAAAGAGAAAAGCTTACCTTCAAGATATTTGTGCA 2094  
 RESULT 5  
 ADM43212  
 ID ADM43212 standard; cDNA; 2094 BP.  
 XX  
 AC ADM43212;  
 XX  
 DT 03-JUN-2004 (first entry)  
 XX  
 DE Human methionine synthase reductase CDS G110A variant.  
 XX  
 KW Human; ss; Methionine synthase reductase polypeptide; HsmTRR; cancer;  
 KW cardiovascular disease; neural tube defect; hyperhomocysteinemia;  
 KW chromosome 5p15.2-p15.3; SNP; single nucleotide polymorphism.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..2094  
 FT /tag= a  
 FT /product= "hsmTRR"  
 FT /partial  
 FT /note= "No stop codon shown"  
 FT variation replace(66,A)  
 FT /tag= b  
 FT /standard\_name= "Single\_nucleotide\_polymorphism"  
 FT replace(110,G)  
 FT /tag= c  
 FT /standard\_name= "Single\_nucleotide\_polymorphism"  
 FT  
 PN US2003082676-A1.  
 XX  
 PD 01-MAY-2003.  
 XX  
 PF 10-AUG-1999; 99US-00371347.  
 XX  
 PR 16-JAN-1998; 98US-0071622P.  
 PR 15-JAN-1999; 99US-00233028.  
 XX  
 PA (GRAV/) GRAVEL R A.  
 PA (ROZE/) ROZEN R.  
 PA (LECL/) LECLERC D.  
 PA (WILS/) WILSON A.  
 PA (ROSE/) ROSENBLATT D.  
 XX  
 PI Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;  
 XX  
 DR WPI; 2003-576610/54.  
 DR P-PSDB; ADM43213.  
 XX  
 PT New substructially pure nucleic acid encoding a mammalian methionine  
 PT synthase reductase polypeptide, useful for diagnosing, preventing or  
 PT treating conditions associated with altered methionine synthase activity,  
 PT e.g. cancer.  
 XX  
 PS Disclosure; SEQ ID NO 43; 26bp; English.  
 XX  
 CC The invention relates to a substructially pure nucleic acid that encodes a  
 CC mammalian methionine synthase reductase polypeptide, HsmTRR, or that  
 CC hybridizes at high stringency to a nucleic acid appearing as ADM43208 or  
 CC ADM43209. Also included are a non-human animal where one or both genetic  
 CC alleles encoding the methionine synthase reductase polypeptide are  
 CC mutated, an antibody that specifically binds the above methionine  
 CC synthase reductase polypeptide, a method of detecting the presence of the  
 CC above polypeptide, a method for detecting sequence variants for  
 CC methionine synthase reductase in a mammal, methods of treating or  
 CC preventing cancer (or cardiovascular disease or neural tube defects) in a  
 CC subject, methods of screening for a compound that modulates methionine  
 CC synthase reductase biological activity and a method for detecting an  
 CC increased risk of developing a neural tube defect in a mammalian embryo  
 CC or foetus. The nucleic acid is useful in diagnosing, preventing or

CC treating conditions associated with altered methionine synthase activity,  
CC such as cancer, cardiovascular disease or neural tube defects, or in  
CC screening for a compound that modulates methionine synthase  
CC biological activity. Naturally occurring variants of the polypeptide are  
CC also associated with hyperhomocysteinaemia. The gene for hMTR is  
CC located on chromosome 5p15.2-p15.3. The present sequence is the coding  
CC sequence of a variant human hMTR cDNA.

XX Sequence 2094 BP; 592 A; 489 C; 480 G; 533 T; 0 U; 0 Other;

Query Match 97.4%; Score 2043; DB 11; Length 2094;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 1; Indels 0; Gaps 0;

Matches 2093; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGAGGTTTCTGTACTATATGCTACACAGCAGGAGACAGGCAAGCCATCGCAGAA 60  
DB 1 ATGAGAGGTTTCTGTACTATATGCTACACAGCAGGAGACAGGCAAGCCATCGCAGAA 60  
QY 61 GAAATGTGTAGCAAGCTGTGTGATCATGATTTTCTGCAATCTTCACTGTATTAGTAA 120  
DB 61 GAAATGTGTAGCAAGCTGTGTGATCATGATTTTCTGCAATCTTCACTGTATTAGTAA 120  
QY 121 TCGGATTAAGTATGACCTTAAGAAACCGAAACAGCTCTCTGTGTGTGTGTGTGTGTG 180  
DB 121 TCGGATTAAGTATGACCTTAAGAAACCGAAACAGCTCTCTGTGTGTGTGTGTGTGTG 180  
QY 181 GGCACCGGAGACCCACCCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAAACCAACA 240  
DB 181 GGCACCGGAGACCCACCCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAAACCAACA 240  
QY 241 CTGCGCGGTGATTTCTTTTGTCTCACCCTGCGGTATGGTTTACTGGGTCTGCGTATTCA 300  
DB 241 CTGCGCGGTGATTTCTTTTGTCTCACCCTGCGGTATGGTTTACTGGGTCTGCGTATTCA 300  
QY 301 TACACCTACTTTTGCATTTGGGGGGGAGATATATGATTAAGSACTTCAAGAGCTTGGAGCC 360  
DB 301 TACACCTACTTTTGCATTTGGGGGGGAGATATATGATTAAGSACTTCAAGAGCTTGGAGCC 360  
QY 361 CGGCAATTTCTATGACACTGTGACATGACATGACTGTGTAGGTTTGAACCTTGTGTGAG 420  
DB 361 CGGCAATTTCTATGACACTGTGACATGACATGACTGTGTAGGTTTGAACCTTGTGTGAG 420  
QY 421 CCGTGAATTTGCTGGACTCTGTGCGCAGCCCTCAAGAAAGATTTTAAAGTCAAGAGAGCAA 480  
DB 421 CCGTGAATTTGCTGGACTCTGTGCGCAGCCCTCAAGAAAGATTTTAAAGTCAAGAGAGCAA 480  
QY 481 GAGGAGTAAGTGTGCGGACCTCCGGTGGCAATCACTGATCTTGAAGACAGACCTTGTG 540  
DB 481 GAGGAGTAAGTGTGCGGACCTCCGGTGGCAATCACTGATCTTGAAGAGACAGACCTTGTG 540  
QY 541 AAGTCAGAGCTGCTACATTTGAATCTCAAGTCAGGCTTCTGAGATTGATTCAGGA 600  
DB 541 AAGTCAGAGCTGCTACATTTGAATCTCAAGTCAGGCTTCTGAGATTGATTCAGGA 600  
QY 601 AGAAAGATTTCTGAGGTTTGAAGCAAAATGCAATGCAAGCAACCAATCCATGTTGTA 660  
DB 601 AGAAAGATTTCTGAGGTTTGAAGCAAAATGCAATGCAAGCAACCAATCCATGTTGTA 660  
QY 661 ATTGAAGACTTTGAGTCTCTACCTTACCGGTTCCGTTACCCCACTCTCAAGAGCTCTCTG 720  
DB 661 ATTGAAGACTTTGAGTCTCTACCTTACCGGTTCCGTTACCCCACTCTCAAGAGCTCTCTG 720  
QY 721 AATATCTCTGTTTACCCCGAATATTTTACAGGTATCATGTGAGAGGTCTCTTGGGCGAG 780  
DB 721 AATATCTCTGTTTACCCCGAATATTTTACAGGTATCATGTGAGAGGTCTCTTGGGCGAG 780  
QY 781 GAGGAAAGCCAGATATCTGTGATTTCAAGAGATCCAGTTTCAAGTGCCTTCAAG 840  
DB 781 GAGGAAAGCCAGATATCTGTGATTTCAAGAGATCCAGTTTCAAGTGCCTTCAAG 840  
QY 841 GCAGTTCAACTTACGTAAGATGACATAAACAACCTGCGGTGATGAAATTGGAATTT 900  
DB 841 GCAGTTCAACTTACGTAAGATGACATAAACAACCTGCGGTGATGAAATTGGAATTT 900

QY 901 TCAATATACAGACTTTTCTATACAGCTTGAGATGCTTACGCGTATCTGCCCTTAACAGT 960  
DB 901 TCAATATACAGACTTTTCTATACAGCTTGAGATGCTTACGCGTATCTGCCCTTAACAGT 960  
QY 961 GATTCTGAGTACAAAGCCTTACCAAGATCTGACCTTGAAGATTAAGAGAGCTGC 1020  
DB 961 GATTCTGAGTACAAAGCCTTACCAAGATCTGACCTTGAAGATTAAGAGAGCTGC 1020  
QY 1021 GTCCCTTTGAAATTAAGGAGACACAAAGAAAGAGCTACTTACCCCGACATATA 1080  
DB 1021 GTCCCTTTGAAATTAAGGAGACACAAAGAAAGAGCTACTTACCCCGACATATA 1080  
QY 1081 CCTGCGGAGATGTTCTCTCCAGTTCAATTTTACCTGTGTGTGAAATTCGAGCAATTCCT 1140  
DB 1081 CCTGCGGAGATGTTCTCTCCAGTTCAATTTTACCTGTGTGTGAAATTCGAGCAATTCCT 1140  
QY 1141 AAAAAGCATTTTGGGAGCCCTTGTGACATACAGATGACAGTGTGAAAAGCCAGG 1200  
DB 1141 AAAAAGCATTTTGGGAGCCCTTGTGACATACAGATGACAGTGTGAAAAGCCAGG 1200  
QY 1201 CTACAGAGCTGTGACATTAACAAAGGAGGAGCCGATTTATAGCCGCTTGTACAGATGCC 1260  
DB 1201 CTACAGAGCTGTGACATTAACAAAGGAGGAGCCGATTTATAGCCGCTTGTACAGATGCC 1260  
QY 1261 TGTGCTGTGTGTGATCTCTCTCTGCTTTCCTTTCGAGGCAACACTCAAGTCTC 1320  
DB 1261 TGTGCTGTGTGTGATCTCTCTCTGCTTTCCTTTCGAGGCAACACTCAAGTCTC 1320  
QY 1321 CTGCTGGAACATTTCTTAACTTCAACCCAGACCAATTCGTGTGAGCTCAAGTTTA 1380  
DB 1321 CTGCTGGAACATTTCTTAACTTCAACCCAGACCAATTCGTGTGAGCTCAAGTTTA 1380  
QY 1381 TTTCAACCCAGGAAGCTCAATTTTGTCTTCAACATTTGGAATTTCTGTCTACCTGACA 1440  
DB 1381 TTTCAACCCAGGAAGCTCAATTTTGTCTTCAACATTTGGAATTTCTGTCTACCTGACA 1440  
QY 1441 ACAGAGGTTCTGCGGAGGAGATGATGACAGCTGTGCTGTGTGTGTGTGTGTGCTTCA 1500  
DB 1441 ACAGAGGTTCTGCGGAGGAGATGATGACAGCTGTGCTGTGTGTGTGTGTGTGCTTCA 1500  
QY 1501 CTTGACGCAAACTTACATGATTCCTCATGAAAGAGGAGGAAACCCCTGCTCTTAAGATA 1560  
DB 1501 CTTGACGCAAACTTACATGATTCCTCATGAAAGAGGAGGAAACCCCTGCTCTTAAGATA 1560  
QY 1561 TCCATCTCTCTGGAACAACAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
DB 1561 TCCATCTCTCTGGAACAACAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
QY 1621 ATAATGTTGGGTCACAGAACCGGACATAGCCCGTTTATTTGGGTTCTTCAACAATAGAG 1680  
DB 1621 ATAATGTTGGGTCACAGAACCGGACATAGCCCGTTTATTTGGGTTCTTCAACAATAGAG 1680  
QY 1681 AAATCTCAAAACAACAACCGAGATGAAATTTTGGAGCAATGTGTGTTTGTGCTGC 1740  
DB 1681 AAATCTCAAAACAACAACCGAGATGAAATTTTGGAGCAATGTGTGTTTGTGCTGC 1740  
QY 1741 AGGCATTAAGATAGGATTAATCTATTCAGAAAAGAGCTCAGACATTTCTTAAGATGGG 1800  
DB 1741 AGGCATTAAGATAGGATTAATCTATTCAGAAAAGAGCTCAGACATTTCTTAAGATGGG 1800  
QY 1801 ATCTTAATCTATTAAGGTTTCTTCTCAAGAGATGCTCTGTGTGGGAGGAGAGGCC 1860  
DB 1801 ATCTTAATCTATTAAGGTTTCTTCTCAAGAGATGCTCTGTGTGGGAGGAGAGGCC 1860  
QY 1861 CAGCAAGATATGTAACAAGCAATCCAGCTTCAATGAGAGAGGAGGAGAAATCTC 1920  
DB 1861 CAGCAAGATATGTAACAAGCAATCCAGCTTCAATGAGAGAGGAGGAGAAATCTC 1920  
QY 1921 CTCAGGAGAACGAGCAATTTTATGTGTGTGAGATGCAAGATATGAGCAAGATGTA 1980  
DB 1921 CTCAGGAGAACGAGCAATTTTATGTGTGTGAGATGCAAGATATGAGCAAGATGTA 1980



QY 1981 CATGATGCCCTTGTGCAATATTAAGCAAGAGTTGGAGTTGAAACTAGAGCAATG 2040  
DB 1981 CATGATGCCCTTGTGCAATATTAAGCAAGAGTTGGAGTTGAAACTAGAGCAATG 2040  
QY 2041 AAAACCTGGCCACTTTAAAGAGAAAAACCTTACCTTCAGATATTGTGCA 2094  
DB 2041 AAAACCTGGCCACTTTAAAGAGAAAAACCTTACCTTCAGATATTGTGCA 2094

RESULT 6  
ADM43209  
ID ADM43209 standard; cDNA; 2094 BP.  
AC ADM43209;  
XX  
XX 03-JUN-2004 (first entry)  
XX  
XX Human methionine synthase reductase CDS G66A variant.  
XX  
XX Human; ss: Methionine synthase reductase polypeptide; HsMTRR; cancer;  
XX cardiovascular disease; neural tube defect; hyperhomocysteinemia;  
XX chromosome 5p15.2-p15.3; SNP; single nucleotide polymorphism.  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX CDS 1..2094  
XX FT /\*tag= a  
XX FT /product= "HsMTRR"  
XX FT /partial  
XX FT /note= "No stop codon shown"  
XX FT replace(66,G)  
XX FT /\*tag= b  
XX FT /standard\_name= "Single\_nucleotide\_polymorphism"  
XX FT replace(110,A)  
XX FT /\*tag= c  
XX FT /standard\_name= "Single\_nucleotide\_polymorphism"  
XX PN US2003082676-A1.  
XX  
XX 01-MAY-2003.  
XX PD  
XX PF 10-AUG-1999; 99US-00371347.  
XX PR 16-JAN-1998; 98US-0071622P.  
XX PR 15-JAN-1999; 99US-00232028.  
XX  
XX PA (GRAV/) GRAVEL R. A.  
XX PA (ROZE/) ROZEN R.  
XX PA (LECL/) LECLEERC D.  
XX PA (WILS/) WILSON A.  
XX PA (ROSE/) ROSENBLATT D.  
XX  
XX PI Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;  
XX WPI: 2003-576610/54.  
XX DR P-PSDB; ADM43211.  
XX  
XX PT New substantially pure nucleic acid encoding a mammalian methionine  
XX PT synthase reductase polypeptide, useful for diagnosing, preventing or  
XX PT treating conditions associated with altered methionine synthase activity,  
XX PT e.g. cancer.  
XX  
XX PS Claim 3; SEQ ID NO 41; 26pp; English.  
XX  
XX The invention relates to a substantially pure nucleic acid that encodes a  
XX mammalian methionine synthase reductase polypeptide, HsMTRR, or that  
XX hybridises at high stringency to a nucleic acid appearing as ADM43208 or  
XX ADM43209. Also included are a non-human animal where one or both genetic  
XX alleles encoding the methionine synthase reductase polypeptide are  
XX mutated, an antibody that specifically binds the above methionine  
XX synthase reductase polypeptide, a method of detecting the presence of the  
XX above polypeptide, a method for detecting sequence variants for

CC methionine synthase reductase in a mammal, methods of treating or  
CC preventing cancer (or cardiovascular disease or neural tube defects) in a  
CC subject, methods of screening for a compound that modulates methionine  
CC synthase reductase biological activity and a method for detecting an  
CC increased risk of developing a neural tube defect in a mammalian embryo  
CC or foetus. The nucleic acid is useful in diagnosing, preventing or  
CC treating conditions associated with altered methionine synthase activity,  
CC such as cancer, cardiovascular disease or neural tube defects, or in  
CC screening for a compound that modulates methionine synthase reductase  
CC biological activity. Naturally occurring variants of the polypeptide are  
CC also associated with hyperhomocysteinemia. The gene for HsMTRR is  
CC located on chromosome 5p15.2-p15.3. The present sequence is the coding  
CC sequence of a variant human HsMTRR cDNA.  
XX  
XX Sequence 2094 BP; 592 A; 489 G; 480 G; 533 T; 0 U; 0 Other;  
XX  
XX Query Match 97.4%; Score 2043; DB 11; Length 2094;  
XX Best Local Similarity 100.0%; Pred. No. 0;  
XX Matches 2093; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ATGAGAGAGTTCTGTTACTATATGCTACACGACGAGGACAGGCAAGGCCATGCCAGAA 60  
DB 1 ATGAGAGAGTTCTGTTACTATATGCTACACGACGAGGACAGGCAAGGCCATGCCAGAA 60  
QY 61 GAAATGTGAGCAAGCTGTGATCATGGAATTTCTGAGATCTTCACTGATTAAGTGA 120  
DB 61 GAAATGTGAGCAAGCTGTGATCATGGAATTTCTGAGATCTTCACTGATTAAGTGA 120  
QY 121 TCCGATTAAGTATGACCTTAAAAACGAAACAGCTCTCTTGTGTTGTGTTCTACACG 180  
DB 121 TCCGATTAAGTATGACCTTAAAAACGAAACAGCTCTCTTGTGTTGTGTTCTACACG 180  
QY 181 GGCACCGAGACCCACCGACACCGCCGCAAGTTGTTAAGAAATACAGAACCAABA 240  
DB 181 GGCACCGAGACCCACCGACACCGCCGCAAGTTGTTAAGAAATACAGAACCAABA 240  
QY 191 GGCACCGAGACCCACCGACACCGCCGCAAGTTGTTAAGAAATACAGAACCAABA 240  
DB 191 GGCACCGAGACCCACCGACACCGCCGCAAGTTGTTAAGAAATACAGAACCAABA 240  
QY 241 CTGCCGTTGATTTCTTGTCTCACTGCGGTATGCGTTACTGCGTCTCGGTATTCAGA 300  
DB 241 CTGCCGTTGATTTCTTGTCTCACTGCGGTATGCGTTACTGCGTCTCGGTATTCAGA 300  
QY 301 TACACCTTACTTTGCAATGGGGGGAAGATATGATTAACGACTTCAAGAGCTTGAGCC 360  
DB 301 TACACCTTACTTTGCAATGGGGGGAAGATATGATTAACGACTTCAAGAGCTTGAGCC 360  
QY 361 CGGCAATTTCTATGACATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 420  
DB 361 CGGCAATTTCTATGACATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 420  
QY 421 CCGTGATTTGCTGAATCTGCGCAGCCCTCAGAAAGATTTTAAAGTCAAGAGAGCA 480  
DB 421 CCGTGATTTGCTGAATCTGCGCAGCCCTCAGAAAGATTTTAAAGTCAAGAGAGCA 480  
QY 481 GAGAGATTAAGTGGGCACTCCCGTGGCATCAGCTGATCTTGAAGACAGACCTTGG 540  
DB 481 GAGAGATTAAGTGGGCACTCCCGTGGCATCAGCTGATCTTGAAGACAGACCTTGG 540  
QY 541 AAGTCAAGCTCTTACATTAATCTCAATGCAATGCAATGCAATGCAATGCAATG 600  
DB 541 AAGTCAAGCTCTTACATTAATCTCAATGCAATGCAATGCAATGCAATGCAATG 600  
QY 601 AAGTCAAGCTCTTACATTAATCTCAATGCAATGCAATGCAATGCAATGCAATG 660  
DB 601 AAGTCAAGCTCTTACATTAATCTCAATGCAATGCAATGCAATGCAATGCAATG 660  
QY 661 ATTGAAGACTTGAAGTCTCACTTACCCGTTGGTATCCCGCACTCTCAAGACCTCTG 720  
DB 661 ATTGAAGACTTGAAGTCTCACTTACCCGTTGGTATCCCGCACTCTCAAGACCTCTG 720  
QY 721 AATATTTCTGTTTACCCCGCAATATTTACAGTACATTCAGAGAGTCTTGGCCAG 780  
DB 721 AATATTTCTGTTTACCCCGCAATATTTACAGTACATTCAGAGAGTCTTGGCCAG 780  
QY 781 GAGGAAGCAGATATCTGTGACTTCAGCAGATCAGTTTCAAGTGCATTTCTTAAG 840



```

Db      781  GAGAAAGCCAGATCTGACTGAGCAGATCCAGTTTCAAGGCCAATTTCAAG
Qy      841  GCAGTCAATCTACGAATGATGCAATAAAACCACTGCGGGAGAAATGGAATT
Db      841  GCAGTCAATCTACGAATGATGCAATAAAACCACTGCGGGAGAAATGGAATT
Qy      901  TCAATACAGACTTTTCTATCAGCCCTGAGATGCTTCAAGCGTGAATCTGCT
Db      901  TCAATACAGACTTTTCTATCAGCCCTGAGATGCTTCAAGCGTGAATCTGCT
Qy      961  GATTCTGAGGTACAAAGCCTACTCTCAAAGCTGAGCTTGAAGATAAAGAGCA
Db      961  GATTCTGAGGTACAAAGCCTACTCTCAAAGCTGAGCTTGAAGATAAAGAGCA
Qy      1021  GTGCTTTGAAAATTAAGGAGACACAAAGAAAGAGACTACCTTACCCGACAT
Db      1021  GTGCTTTGAAAATTAAGGAGACACAAAGAAAGAGACTACCTTACCCGACAT
Qy      1081  CCTGCGGAGATGTTCTCTCAAGTTCATTTTACCTGCTGATGCTGAAATCC
Db      1081  CCTGCGGAGATGTTCTCTCAAGTTCATTTTACCTGCTGATGCTGAAATCC
Qy      1141  AAAAAGGCAATTTTTCGAGCCCTTGTGACTATACAGTGAAGTCTGAAAAG
Db      1141  AAAAAGGCAATTTTTCGAGCCCTTGTGACTATACAGTGAAGTCTGAAAAG
Qy      1201  CTACAGAGCTGTGCACTAAACAAAGGGAGCGGATATAGCCGCTTGTACG
Db      1201  CTACAGAGCTGTGCACTAAACAAAGGGAGCGGATATAGCCGCTTGTACG
Qy      1261  TGTGCTGCTGTGTGATCTCTCTGCTGCTTCCCTTCTTGCCAGCCACCTC
Db      1261  TGTGCTGCTGTGTGATCTCTCTGCTGCTTCCCTTCTTGCCAGCCACCTC
Qy      1321  CTGCTGCAATCTTCTTAACTTCAACCCAGACATTTGCTGCAAGTCTCA
Db      1321  CTGCTGCAATCTTCTTAACTTCAACCCAGACATTTGCTGCAAGTCTCA
Qy      1381  TTTCAACCCAGAAAGCTCAATTTTGTCTTCAACATTTGCTGCAAGTCTC
Db      1381  TTTCAACCCAGAAAGCTCAATTTTGTCTTCAACATTTGCTGCAAGTCTC
Qy      1441  ACAAGAGTCTGCGGAGGAGATATGTAAGAGCTGCGCTTGTGCTTCA
Db      1441  ACAAGAGTCTGCGGAGGAGATATGTAAGAGCTGCGCTTGTGCTTCA
Qy      1501  CTTCAAGCCAAACATATGATCCCATGAAAGACGCGGAAAGCCCTGCTC
Db      1501  CTTCAAGCCAAACATATGATCCCATGAAAGACGCGGAAAGCCCTGCTC
Qy      1561  TCCATCTCTCTCGAACAACAAATTTCTTCACTTACAGATGATGCTTCA
Db      1561  TCCATCTCTCTCGAACAACAAATTTCTTCACTTACAGATGATGCTTCA
Qy      1621  ATATATGTTGGTCCAGGACCGGATAGCCCGTTTATGTTGCTTCAACAT
Db      1621  ATATATGTTGGTCCAGGACCGGATAGCCCGTTTATGTTGCTTCAACAT
Qy      1681  AAATCTCCAAAGAACACCCAGATGAAATTTTGAGCAATGTTGTTTGG
Db      1681  AAATCTCCAAAGAACACCCAGATGAAATTTTGAGCAATGTTGTTTGG
Qy      1741  AGGATATAGATAGGATATATCTATTCAGAAAGAGCTCAGACATTTCT
Db      1741  AGGATATAGATAGGATATATCTATTCAGAAAGAGCTCAGACATTTCT
Qy      1801  ATCTTAATCTATCTAAAGGTTTCTTCTCAAGAGATGCTCTGTTGGG
Db      1801  ATCTTAATCTATCTAAAGGTTTCTTCTCAAGAGATGCTCTGTTGGG
Qy      1861  CCAGCAAGTATGTAAGAGACATCAAGTTCAAGCCGACGAGTGGGAGAA

```

```

Db      1861  CCAGCAAGTATGTAAGAGACATCAAGTTCAAGCCGACGAGTGGGAGAA
Qy      1921  CTCACAGAGAACGGCCATATTTATGTGTGAGATGCAAGAAATATG
Db      1921  CTCACAGAGAACGGCCATATTTATGTGTGTGAGATGCAAGAAATATG
Qy      1981  CATGATGCCCTTGTGCAATATATACGAAAGAGTTGAGTTGAAAA
Db      1981  CATGATGCCCTTGTGCAATATATACGAAAGAGTTGAGTTGAAAA
Qy      2041  AAAACCTGCGCACTTTAAAGAGAAAGAGTACCTCAGGATATTTG
Db      2041  AAAACCTGCGCACTTTAAAGAGAAAGAGTACCTCAGGATATTTG

RESULT 7
AAAS8935
ID  AAAS8935 standard; DNA: 3259 BP.
XX
AC  AAAS8935;
XX
DT  07-NOV-2000 (first entry)
XX
DE  DNA encoding a human methionine synthase reductase polypeptide.
XX
KW  Human; methionine synthase reductase; MTRR; cancer;
KM  cardiovascular disease; Down's Syndrome; neural tube defect;
XX  premature coronary artery disease; ss.
XX
OS  Homo sapiens.
XX
FH  Key Location/Qualifiers
FT  CDS 80..2176
FT  /tag= a
FT  /product= "methionine synthase reductase"
XX
FN  WC2000042196-A2.
XX
PD  20-JUL-2000.
XX
PF  14-JAN-2000; 2000MO-IB000209.
XX
PR  15-JAN-1999; 99US-00232028.
XX  10-AUG-1999; 99US-00371347.
XX
PA  (UWMC-) UNIV MCGILL.
XX
PI  Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;
DR  MPI; 2000-466131/40.
XX  P-PSDB; AAB07591.
XX
PT  Mammalian methionine synthase reductase nucleic acid used for detecting
XX  an increased risk of developing a neural tube defect. Down's Syndrome or
XX  cardiovascular disease in a mammalian embryo or fetus.
XX
PS  Claim 3; Fig 3; 85bp; English.
XX
CC  The present sequence encodes a human methionine synthase reductase (MTRR)
CC  polypeptide. Inhibitors of MTRR polypeptide and polynucleotide are used
CC  for treating or preventing cancer, cardiovascular disease, Down's
CC  Syndrome or neural tube defects in a subject. The cardiovascular disease
CC  is premature coronary artery disease. The compounds are detected by
CC  methods which screen for modulators of MTRR biological activity. MTRR
CC  polypeptide or nucleic acid is examined for the presence of a
CC  polymorphism in the parents or the embryo or foetus, and the information
CC  used for detecting an increased risk of an embryo or foetus developing
CC  cancer, cardiovascular disease, Down's Syndrome or neural tube defects
XX
SQ  Sequence 3259 BP; 944 A; 706 C; 663 G; 946 T; 0 U; 0 Other;
Query Match 95.1%; Score 1995; DB 3; Length 3259;
Best Local Similarity 99.9%; Pred. No. 0;

```

Matches 2095; Conservative 0; Mismatches 2; Indels 0; Gaps 0;				
QY	1	ATGAGAGGTTCTGTACTATATGTACACAGACGAGGACAGGCAAAAGGCATCGAGAA	60	
DB	80	ATGAGAGGTTCTGTACTATATGTACACAGACGAGGACAGGCAAAAGGCATCGAGAA	139	
QY	61	GAATGTGTGACAGAGCTGTGTACATGTATTTCTGCAGATCTTCACTGATTTAGTGA	120	
DB	140	GAATGTGTGACAGAGCTGTGTACATGTATTTCTGCAGATCTTCACTGATTTAGTGA	199	
QY	121	TCCGATAGTATGACTTAAACCGAAGAGCTCTTGTGTGTGTGTGTCTACACG	180	
DB	200	TCCGATAGTATGACTTAAACCGAAGAGCTCTTGTGTGTGTGTGTCTACACG	259	
QY	181	GGACACCGGAGACCCGACACAGCCGCAAGTTGTATAGGAATATAGAAACCAACA	240	
DB	260	GGACACCGGAGACCCGACACAGCCGCAAGTTGTATAGGAATATAGAAACCAACA	319	
QY	241	CTGCCGTTGATTTCTTGTCTCACTGCGGTATGAGTTAGGTTCTCGGTGATTCAGAA	300	
DB	320	CTGCCGTTGATTTCTTGTCTCACTGCGGTATGAGTTAGGTTCTCGGTGATTCAGAA	379	
QY	301	TACACCTACTTTTGCATGAGGAGGAGATATATTAACGACTTCAAGAGCTTGAGCC	360	
DB	380	TACACCTACTTTTGCATGAGGAGGAGATATATTAACGACTTCAAGAGCTTGAGCC	439	
QY	361	CGGCATTTCTATGACCTGACATGTAGATGTGTGTGTGTATTAACCTTGTGTGTGAG	420	
DB	440	CGGCATTTCTATGACCTGACATGTAGATGTGTGTGTGTATTAACCTTGTGTGTGAG	499	
QY	421	CCGTGATTTGTGACCTTGTGCGACGCTCAGAAACATTTTATAGTCAAGACAGAGCAA	480	
DB	500	CCGTGATTTGTGACCTTGTGCGACGCTCAGAAACATTTTATAGTCAAGACAGAGCAA	559	
QY	481	GAGGAGATAGTGGCGACCTCCGGTGGCATCACTGCATCTTGAAGA CAGACCTTGTG	540	
DB	560	GAGGAGATAGTGGCGACCTCCGGTGGCATCACTGCATCTTGAAGA CAGACCTTGTG	619	
QY	541	AAGTCAGACCTGTACATTTGAATCTCAAGTCGACCTTGAATTCATGATTCAGAA	600	
DB	620	AAGTCAGACCTGTACATTTGAATCTCAAGTCGACCTTGAATTCATGATTCAGAA	679	
QY	601	AGAAAGATTTGAGGTTTGAAGCAAAATGACAGTGAACGACCAATCCATGTTGTA	660	
DB	680	AGAAAGATTTGAGGTTTGAAGCAAAATGACAGTGAACGACCAATCCATGTTGTA	739	
QY	661	ATTGAAGACTTGAAGTCTCACTTACCCGTTCCGTAACCCCACTCTCAAGCCTCTCTG	720	
DB	740	ATTGAAGACTTGAAGTCTCACTTACCCGTTCCGTAACCCCACTCTCAAGCCTCTCTG	799	
QY	721	AATATTTCTGTGTTTACCCCAAGATTTTACAGGTACATCTGAGAGAGTCTTGGCCAG	780	
DB	800	AATATTTCTGTGTTTACCCCAAGATTTTACAGGTACATCTGAGAGAGTCTTGGCCAG	859	
QY	781	GAGGAAAGCAAGTATCTGTGACTTCAAGAGATCCAGTTTTCAGGTGCAATTTCAAA	840	
DB	860	GAGGAAAGCAAGTATCTGTGACTTCAAGAGATCCAGTTTTCAGGTGCAATTTCAAA	919	
QY	841	GCAATTCACTTACATGATGATGCAATTAACCACTCTGTGTGTAGAAATTTGACATT	900	
DB	920	GCAATTCACTTACATGATGATGCAATTAACCACTCTGTGTGTAGAAATTTGACATT	979	
QY	901	TCAAAATCAGACTTTTCTATGAGCTGTGAGATGCTTCAAGGTATCTGCTTAAACAG	960	
DB	980	TCAAAATCAGACTTTTCTATGAGCTGTGAGATGCTTCAAGGTATCTGCTTAAACAG	1039	
QY	961	GATTTGAGGTACAAAGCTTCAAAAGCTGACCTTGAAGATTAAGAGAGACCTGC	1020	
DB	1040	GATTTGAGGTACAAAGCTTCAAAAGCTGACCTTGAAGATTAAGAGAGACCTGC	1099	
QY	1021	GTCCTTTTGAATAAAGGACACACAAAGAAAGAGCTTACCTTACCCAGCATATA	1080	
DB	1100	GTCCTTTTGAATAAAGGACACACAAAGAAAGAGCTTACCTTACCCAGCATATA	1159	

QY	1081	CTGCGGAGATGTTCTCTCCAGTTCAATTTTACCTGTGTCTTGAATCCGACCAATTCCT	1140	
DB	1160	CTGCGGAGATGTTCTCTCCAGTTCAATTTTACCTGTGTCTTGAATCCGACCAATTCCT	1219	
QY	1141	AAAAAGCAATTTTTCGAGCCCTTGTGACATATACAGTACAGTGTCTGAAAAGCCAG	1200	
DB	1220	AAAAAGCAATTTTTCGAGCCCTTGTGACATATACAGTACAGTGTCTGAAAAGCCAG	1279	
QY	1201	CTACAGAGCTGTGACATTAAGGGGACCGCATATATAGCCGCTTGTAGAGATGCG	1260	
DB	1280	CTACAGAGCTGTGACATTAAGGGGACCGCATATATAGCCGCTTGTAGAGATGCG	1339	
QY	1261	TGTGCTGTGTGTGATCTCTCTGCTTCCCTTCTTGCAGACCACTCACTGCTC	1320	
DB	1340	TGTGCTGTGTGTGATCTCTCTGCTTCCCTTCTTGCAGACCACTCACTGCTC	1399	
QY	1321	GTGCTGAAATCTTCTTAACTTGAACCCAGACATATGTGTGCAAGCTCAAGTTA	1380	
DB	1400	GTGCTGAAATCTTCTTAACTTGAACCCAGACATATGTGTGCAAGCTCAAGTTA	1459	
QY	1381	TTTCAACCCAGAAAGCTCATTGTTCTTCAATTTGAGAAATTTCTGTCTACGCA	1440	
DB	1460	TTTCAACCCAGAAAGCTCATTGTTCTTCAATTTGAGAAATTTCTGTCTACGCA	1519	
QY	1441	ACAGAGTTCTGCGAAGGAGATATGTACAGGCTGGCTGTGTGTGCTTCAATT	1500	
DB	1520	ACAGAGTTCTGCGAAGGAGATATGTACAGGCTGGCTGTGTGTGCTTCAATT	1579	
QY	1501	CTTCAACCAACATATGATCCATGAGAGACAGCGGAAAAGCCCTGTGCTCTTAAGATA	1560	
DB	1580	CTTCAACCAACATATGATCCATGAGAGACAGCGGAAAAGCCCTGTGCTCTTAAGATA	1639	
QY	1561	TTCATCTCTCTGAAACCAATTTCTTCACTTACAGATGACCCCTCAATCCCAATC	1620	
DB	1640	TTCATCTCTCTGAAACCAATTTCTTCACTTACAGATGACCCCTCAATCCCAATC	1699	
QY	1621	ATATATGTTGGTTCAGAAACCGGCAATAGCCCGTTTATGGGTTCTTAACAATAGAG	1680	
DB	1700	ATATATGTTGGTTCAGAAACCGGCAATAGCCCGTTTATGGGTTCTTAACAATAGAG	1759	
QY	1681	AAACTCCAAAGAACACCCAGATGGAATTTTGAAGCATGTGTGTTTTTGTGCTGC	1740	
DB	1760	AAACTCCAAAGAACACCCAGATGGAATTTTGAAGCATGTGTGTTTTTGTGCTGC	1819	
QY	1741	AGGCATAGGATAGGATATATCTATTCAGAAAAGAGCTCAGACATTTCTTAAGCATGG	1800	
DB	1820	AGGCATAGGATAGGATATATCTATTCAGAAAAGAGCTCAGACATTTCTTAAGCATGG	1879	
QY	1801	ATCTTAATCATCTAAGGTTCTTCTCAAGAGATGCTCTGTGGGAGAGGAAAGCC	1860	
DB	1880	ATCTTAATCATCTAAGGTTCTTCTCAAGAGATGCTCTGTGGGAGAGGAAAGCC	1939	
QY	1861	CCAGCAAGTATGTACAGACAAATCCAGCTTCAATGACAGAGGTGCGAATCTTC	1920	
DB	1940	CCAGCAAGTATGTACAGACAAATCCAGCTTCAATGACAGAGGTGCGAATCTTC	1999	
QY	1921	CTTCAGAGAAAGGCGCATATTTATGTGTGTGAGATGCAAAAGATTTGGCCAAAGTGA	1980	
DB	2000	CTTCAGAGAAAGGCGCATATTTATGTGTGTGAGATGCAAAAGATTTGGCCAAAGTGA	2059	
QY	1981	CATGATGACCTTGTGCAATTAATTAAGCAAGGTTGAGTTGAAAACCTAAGAGCATG	2040	
DB	2060	CATGATGACCTTGTGCAATTAATTAAGCAAGGTTGAGTTGAAAACCTAAGAGCATG	2119	
QY	2041	AAAACCTGTGCACTTAAAGAAAGAAACGCTACCTTCAAGATATTTGTCTATA	2097	
DB	2120	AAAACCTGTGCACTTAAAGAAAGAAACGCTACCTTCAAGATATTTGTCTATA	2176	

RESULT 8  
AD087538  
ID AD087538 standard; cDNA; 3270 BP.

XX ADQ87538;  
XX 07-OCT-2004 (first entry)  
XX  
XX Human tumour-associated antigenic target (TAT) cDNA sequence #4416.  
XX  
XX human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;  
XX cancer; cell proliferative disorder; gene; ss.  
XX  
XX Homo sapiens.  
XX  
XX NO2004060270-A2.  
XX  
XX 22-JUL-2004.  
XX  
XX 15-OCT-2003; 2003WO-US029126.  
XX  
XX 18-OCT-2002; 2002US-0418988P.  
XX  
XX (GENTH ) GENENTECH INC.  
XX (WUTD/) WU T D.  
XX (ZHOU/) ZHOU Y.  
XX  
XX Wu TD, Zhou Y;  
XX  
XX WPI; 2004-534300/51.  
XX  
XX New nucleic acid molecule and encoded polypeptide, for diagnosing,  
XX preventing or treating cell proliferative disorders such as cancer.  
XX  
XX Claim 1; SEQ ID NO 4416; 5504dp; English.  
XX  
XX The present invention describes an isolated tumour-associated antigenic  
XX target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide  
XX sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of  
XX (a); (c) the complement of (a) or (b); (d) a sequence that has 80%  
XX sequence identity to (a)-(c); or (e) a sequence that hybridises to (a) -  
XX (c). Also described: (1) an expression vector comprising the above  
XX nucleic acid; (2) a host cell comprising the above expression vector; (3)  
XX a process for producing a polypeptide; (4) an isolated polypeptide  
XX comprising: (a) an amino acid sequence encoded by any of the above  
XX nucleotide sequences; (b) an amino acid sequence encoded by the full-  
XX length coding region of the above nucleotide sequences; or (c) a sequence  
XX having at least 80% identical to (a) or (b); (5) a chimeric polypeptide  
XX comprising the above polypeptide fused to a heterologous polypeptide; (6)  
XX an isolated antibody that binds to the above polypeptide; (7) a process  
XX for producing the antibody; (8) an isolated oligopeptide that binds to  
XX the above polypeptide; (9) a tumour-associated antigenic target (TAT)  
XX binding organic molecule that binds to the above polypeptide; (10) a  
XX composition of matter comprising the above (chimeric) polypeptide,  
XX antibody, oligopeptide or TAT binding organic molecule, in combination  
XX with a carrier; (11) an article of manufacture comprising a container and  
XX the composition of matter contained within the container; (12) methods of  
XX inhibiting the growth of a cell that expresses the above protein, where  
XX the growth of the cell is at least in part dependent upon a growth  
XX potentialising effect of the above protein; (13) a method of  
XX therapeutically treating a mammal having a cancerous tumour comprising  
XX cells that express the above protein; (14) a method of determining the  
XX presence of a protein in a sample suspected of containing the protein  
XX described above; (15) methods of diagnosing the presence of a tumour in a  
XX mammal; (16) a method for treating or preventing a cell proliferative  
XX disorder associated with increased expression or activity of the above  
XX protein; and (17) a method of binding an antibody, oligopeptide or  
XX organic molecule to a cell that expresses the protein described above.  
XX The TAT sequences have cytostatic activities, and can be used in gene  
XX therapy. The composition and methods are useful for diagnosing,  
XX preventing or treating cancer. The composition is also used for preparing  
XX a medicament for the therapeutic treatment or diagnostic detection of a  
XX cell proliferative disorder or cancer. The present sequence represents a  
XX human TAT cDNA sequence from the present invention.  
XX  
XX Sequence 3270 BP; 934 A; 702 C; 680 G; 954 T; 0 U; 0 Other;

Query Match 92.7%; Score 1944; DB 13; Length 3270;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2094; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 ATGAGAGGTTTCTGTACTATATGCTACACAGCAGGACAGGCAAGCCATCGAGAA 60  
DB 112 ATGAGAGGTTTCTGTACTATATGCTACACAGCAGGACAGGCAAGCCATCGAGAA 171  
QY 61 GAAATGTGAGCAAGCTGTGTACATGATGATTTCTGAGATCTTCATGATTAAGTAA 120  
DB 172 GAAATGTGAGCAAGCTGTGTACATGATGATTTCTGAGATCTTCATGATTAAGTAA 231  
QY 121 TCCGATTAAGTAACTTAAACCGAAGAGCTCTCTGTTGTTGTTTACACAG 180  
DB 232 TCCGATTAAGTAACTTAAACCGAAGAGCTCTCTGTTGTTGTTTACACAG 291  
QY 181 GGCACCGGAGACCCAGCCGACAGCCGAGTTTGAAGAAATPACAGAACCAACA 240  
DB 292 GGCACCGGAGACCCAGCCGACAGCCGAGTTTGAAGAAATPACAGAACCAACA 351  
QY 241 CTGCGGTTGATTTCTTGTCTCACTGCGGTATGAGTTAGGATCTCGGTATGAGAA 300  
DB 352 CTGCGGTTGATTTCTTGTCTCACTGCGGTATGAGTTAGGATCTCGGTATGAGAA 411  
QY 301 TACACCTACTTTTGAATGGGGGAAATATATGATTAACGACTTCAAGCTTGAACC 360  
DB 412 TACACCTACTTTTGAATGGGGGAAATATATGATTAACGACTTCAAGCTTGAAGCC 471  
QY 361 CGGATTTCTATGACACTGACATGAGATGATGATGATGATTAACCTTGGTGGTTAG 420  
DB 472 CGGATTTCTATGACACTGACATGAGATGATGATGATGATTAACCTTGGTGGTTAG 531  
QY 421 CCGTGAATGCTGAGCTGTGCGCAGCCCTCAGAAACATTTTATGATCAAGAGACAA 480  
DB 532 CCGTGAATGCTGAGCTGTGCGCAGCCCTCAGAAACATTTTATGATCAAGAGACAA 591  
QY 481 GAGGAGATTAAGTGGCGCACTCCGCTGGCATCACTGATCTTGAAGACAGACCTTGTG 540  
DB 592 GAGGAGATTAAGTGGCGCACTCCGCTGGCATCACTGATCTTGAAGACAGACCTTGTG 651  
QY 541 AACTGAGCTGTACACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
DB 652 AACTGAGCTGTACACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 711  
QY 601 AGAAGAGATTCAGGTTTGAAGCAAAATGACAGTAAACAGCAACCAATCAATGTTGTA 660  
DB 712 AGAAGAGATTCAGGTTTGAAGCAAAATGACAGTAAACAGCAACCAATCAATGTTGTA 771  
QY 661 ATTGAAGCTTTGATGCTCTCACTTACCCGTTGGGTACCCCACTCTCAAGCTCTCTG 720  
DB 772 ATTGAAGCTTTGATGCTCTCACTTACCCGTTGGGTACCCCACTCTCAAGCTCTCTG 831  
QY 721 AATATTCGTTTACCCCGAGAAATTTTACAGTATCATGACAGAGCTCTTGGCCAG 780  
DB 832 AATATTCGTTTACCCCGAGAAATTTTACAGTATCATGACAGAGCTCTTGGCCAG 891  
QY 781 GAGGAAAGCAATATCTGTGACTTCAAGCATTCAGTTTTCAGTGGCCAAATTTCAAG 840  
DB 892 GAGGAAAGCAATATCTGTGACTTCAAGCATTCAGTTTTCAGTGGCCAAATTTCAAG 951  
QY 841 GCAATTCACCTTCACTAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
DB 952 GCAATTCACCTTCACTAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1011  
QY 901 TCAATTAACAGCTTTTCTTCAAGCTGAGAGAGCTTCAAGCTGAGAGAGCTTCAAGT 960  
DB 1012 TCAATTAACAGCTTTTCTTCAAGCTGAGAGAGCTTCAAGCTGAGAGAGCTTCAAGT 1071  
QY 961 GATTTCAGGTTACAAAGCTCAATCAAGAGCTCAAGCTCAAGTCAAGTCAAGTCAAGT 1020  
DB 1072 GATTTCAGGTTACAAAGCTCAATCAAGAGCTCAAGCTCAAGTCAAGTCAAGTCAAGT 1131

1021 GTCTTTGAAATAAAGGAGACACAAAGAAAGAGCTACCTTACCCAGCATATA 1080  
 1132 GTCTTTTGAATAAAGGAGACACAAAGAAAGAGCTACCTTACCCAGCATATA 1191  
 1081 CCTGCGGAGATGTTCTCTCCAGTTCAATTTTACCTGCTGTTGAAATCCGAGCAATTCCT 1140  
 1192 CCTGCGGAGATGTTCTCTCCAGTTCAATTTTACCTGCTGTTGAAATCCGAGCAATTCCT 1251  
 1141 AAAAAGGCAATTTTGGAGAGCCCTTGGAGCTATACAGAGAGCTGTAAGAAAGCCAGG 1200  
 1252 AAAAAGGCAATTTTGGAGAGCCCTTGGAGCTATACAGAGAGCTGTAAGAAAGCCAGG 1311  
 1201 CTACAGAGCTGTGACGTAAACAAGGGGAGCCGATTAATAGCCGCTTGTACAGAGTGC 1260  
 1312 CTACAGAGCTGTGACGTAAACAAGGGGAGCCGATTAATAGCCGCTTGTACAGAGTGC 1371  
 1261 TGTGCTGCTGTTGGATCTCTCTGCTTCCCTTCTTGGCAAGCAGCACTCACTCTC 1320  
 1372 TGTGCTGCTGTTGGATCTCTCTGCTTCCCTTCTTGGCAAGCAGCACTCACTCTC 1431  
 1321 CTGCTGGAACATCTTCTAACTTCAACCCAGACATATTCGTCAGAGCTCAAGTTTA 1380  
 1432 CTGCTGGAACATCTTCTAACTTCAACCCAGACATATTCGTCAGAGCTCAAGTTTA 1491  
 1381 TTTCACCCAGAAAGCTCCTATTTGTCTTCAACATTTGTGGAATTTCTGTCTACCTGCACA 1440  
 1492 TTTCACCCAGAAAGCTCCTATTTGTCTTCAACATTTGTGGAATTTCTGTCTACCTGCACA 1551  
 1441 ACAGAGGTTCTGCGAAGGAGATGTATACAGGCTGCGCTGCTGTTGTTGCTTCAAGTT 1500  
 1552 ACAGAGGTTCTGCGAAGGAGATGTATACAGGCTGCGCTGCTGTTGTTGCTTCAAGTT 1611  
 1501 CTTCAGCCCAACATACATGATCTCCATGAAAGACAGGGGAAAGCCCTGCTCTTAAGATA 1560  
 1612 CTTCAGCCCAACATACATGATCTCCATGAAAGACAGGGGAAAGCCCTGCTCTTAAGATA 1671  
 1561 TCCATCTCTCTCTGGAACAACAATTTCTTCCATCTACAGATACCCCTCAATCCCATC 1620  
 1672 TCCATCTCTCTCTGGAACAACAATTTCTTCCATCTACAGATACCCCTCAATCCCATC 1731  
 1621 ATATATGTTGGGTCAGAGAACCCGACATAGCCCTTATTTGGGTTCTCTACAACTAGAGAG 1680  
 1732 ATATATGTTGGGTCAGAGAACCCGACATAGCCCTTATTTGGGTTCTCTACAACTAGAGAG 1791  
 1681 AAATCTCAAGAAACAACCCAGATGGAATTTTGGAGCAATGTGTTGTTTGGCTGC 1740  
 1792 AAATCTCAAGAAACAACCCAGATGGAATTTTGGAGCAATGTGTTGTTTGGCTGC 1851  
 1741 AGGCAATAGGATAGGATTAATCTATTCAAGAAAGAGCTCAGACATTTCTTAAGCATGGG 1800  
 1852 AGGCAATAGGATAGGATTAATCTATTCAAGAAAGAGCTCAGACATTTCTTAAGCATGGG 1911  
 1801 ATCTTAATCTATTAAGGTTTCTTCTCAAGAGATGCTCTGTTGGGAGAGAGAGGCC 1860  
 1912 ATCTTAATCTATTAAGGTTTCTTCTCAAGAGATGCTCTGTTGGGAGAGAGAGGCC 1971  
 1861 CCAGCAAGATATGTATCAAGACAATCCAGCTTCATGCGCAGACAGAGTGGCGAAGATCTTC 1920  
 1972 CCAGCAAGATATGTATGTATGTATGTATGTATGTATGTATGTATGTATGTATGTATGTAT 2031  
 1921 CTTCAGAGAGAGAGGAGGATTTATGTATGTATGTATGTATGTATGTATGTATGTATGTAT 1980  
 2032 CTTCAGAGAGAGAGGAGGATTTATGTATGTATGTATGTATGTATGTATGTATGTATGTAT 2091  
 1981 CATGATGCTCTGTGTGCAAT 2040  
 2092 CATGATGCTCTGTGTGCAAT 2151  
 2041 AAAACCTGGCCACTTTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 2097  
 2152 AAAACCTGGCCACTTTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 2208

RESULT 9  
 ID ADM43216  
 AC ADM43216; standard; cDNA; 2091 BP.  
 DT 03-JUN-2004 (first entry)  
 DE Human methionine synthase reductase CDS del 1726-1728 variant.  
 KW Human; ss; Methionine synthase reductase polypeptide; HsMTRR; cancer;  
 KW cardiovascular disease; neural tube defect; hyperhomocysteinemia;  
 KW chromosome 5p15.2-p15.3; SNP; single nucleotide polymorphism.  
 OS Homo sapiens.  
 XX  
 FH Location/Qualifiers  
 FT 1..2091  
 FT /tag= a  
 FT /product= "HsMTRRdelR59"  
 FT /partial  
 FT /note= "No stop codon shown"  
 FT replace (66..A)  
 FT /tag= b  
 FT /standard\_name= "Single\_nucleotide\_polymorphism"  
 FT replace (110..A)  
 FT /tag= c  
 FT /standard\_name= "Single\_nucleotide\_polymorphism"  
 FT replace (1726..TTGT)  
 FT /tag= d  
 PN US2003082676-A1.  
 XX  
 PD 01-MAY-2003.  
 XX  
 PF 10-AUG-1999; 99US-00371347.  
 XX  
 PR 16-JAN-1998; 98US-0071622P.  
 PR 15-JUN-1999; 99US-00232028.  
 XX  
 PA (GRAV/) GRAVEL R A.  
 PA (ROZE/) ROZEN R.  
 PA (LECL/) LECLERC D.  
 PA (WILS/) WILSON A.  
 PA (ROSE/) ROSENBLATT D.  
 XX  
 PI Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;  
 XX  
 DR WPI; 2003-576610/54.  
 DR P-PDB; ADM43217.  
 XX  
 PT New substantially pure nucleic acid encoding a mammalian methionine  
 PT synthase reductase polypeptide, useful for diagnosing, preventing or  
 PT treating conditions associated with altered methionine synthase activity,  
 PT e.g. cancer.  
 XX  
 PS Disclosure; SEQ ID NO 45; 26pp; English.  
 XX  
 CC The invention relates to a substantially pure nucleic acid that encodes a  
 CC mammalian methionine synthase reductase polypeptide, HsMTRR, or that  
 CC hybridizes at high stringency to a nucleic acid appearing as ADM43208 or  
 CC ADM43209. Also included are a non-human animal where one or both genetic  
 CC alleles encoding the methionine synthase reductase polypeptide are  
 CC mutated, an antibody that specifically binds the above methionine  
 CC synthase reductase polypeptide, a method of detecting the presence of the  
 CC above polypeptide, a method for detecting sequence variants for  
 CC methionine synthase reductase in a mammal, methods of treating or  
 CC preventing cancer (or cardiovascular disease or neural tube defects) in a  
 CC subject, methods of screening for a compound that modulates methionine  
 CC synthase reductase biological activity and a method for detecting an  
 CC increased risk of developing a neural tube defect in a mammalian embryo  
 CC or foetus. The nucleic acid is useful in diagnosing, preventing or  
 CC treating conditions associated with altered methionine synthase activity,

CC such as cancer, cardiovascular disease or neural tube defects, or in  
CC screening for a compound that modulates methionine synthase reductase  
CC biological activity. Naturally occurring variants of the polypeptide are  
CC also associated with hyperhomocysteinemia. The gene for HmTR is  
CC located on chromosome 5p15.2-p15.3. The present sequence is the coding  
CC sequence of a variant human hmTRR cDNA.

XX Sequence 2091 BP; 591 A; 489 C; 480 G; 531 T; 0 U; 0 Other;

Query Match 88.3%; Score 1851; DB 11; Length 2091;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2091; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 ATGAGAGGTTTCTGTTACTATATGCTACACAGCAGGACAGCAAGCCATCGCAGAA 60  
DB 1 ATGAGAGGTTTCTGTTACTATATGCTACACAGCAGGACAGCAAGCCATCGCAGAA 60  
QY 61 GAAATGTGTAGCAAGCTGTGTATGATGATGATTTCTGACATCTTCACTGTATTAAGGAA 120  
DB 61 GAAATGTGTAGCAAGCTGTGTATGATGATGATTTCTGACATCTTCACTGTATTAAGGAA 120  
QY 121 TCCGATTAAGTATGACTTAAACCGAAACAGCTCTCTGTGTGTGTGTCTTACCCAG 180  
DB 121 TCCGATTAAGTATGACTTAAACCGAAACAGCTCTCTGTGTGTGTGTCTTACCCAG 180  
QY 181 GGCACCCGGAACCAACCCGACACAGCCCGCAAGTTTGTAAAGAAATACGAACCAACA 240  
DB 181 GGCACCCGGAACCAACCCGACACAGCCCGCAAGTTTGTAAAGAAATACGAACCAACA 240  
QY 241 CTGCGGTTGATTTCTTGTGCACCTGCGGTAAGGTTTACCTGGGTCGGGTAATCGAA 300  
DB 241 CTGCGGTTGATTTCTTGTGCACCTGCGGTAAGGTTTACCTGGGTCGGGTAATCGAA 300  
QY 301 TACACCTACTTTTGGCAATGGGGGGAAGTAATTGATTAAGCACTTCAAGAGCTTGAAGC 360  
DB 301 TACACCTACTTTTGGCAATGGGGGGAAGTAATTGATTAAGCACTTCAAGAGCTTGAAGC 360  
QY 361 CGGCAATTTCTATGACACTGACATGACATGACGTGTAGGTTTGAACCTTGTGTGAG 420  
DB 361 CGGCAATTTCTATGACACTGACATGACATGACGTGTAGGTTTGAACCTTGTGTGAG 420  
QY 421 CCGGTGATGTGTGACCTGTGGCCAGCCCTCAGAAAGATTTTAAAGTAACAGAGGACAA 480  
DB 421 CCGGTGATGTGTGACCTGTGGCCAGCCCTCAGAAAGATTTTAAAGTAACAGAGGACAA 480  
QY 481 GAGGAGATTAAGTGTGCGACTCCCGGTGACATCACTGCACTCTTGAAGACAGACTTGTG 540  
DB 481 GAGGAGATTAAGTGTGCGACTCCCGGTGACATCACTGCACTCTTGAAGACAGACTTGTG 540  
QY 541 AAGTCAAGCTGTACACATTTGAATCTCAAGTCAAGCTTTCAGATTTCAGGA 600  
DB 541 AAGTCAAGCTGTACACATTTGAATCTCAAGTCAAGCTTTCAGATTTCAGGA 600  
QY 601 AGAAAGATTTCTGAGTTTGAAGCAAAATGCAATGCAAGCAACCAATCAATGTTGTA 660  
DB 601 AGAAAGATTTCTGAGTTTGAAGCAAAATGCAATGCAAGCAACCAATCAATGTTGTA 660  
QY 661 ATTGAAGACTTGAAGTCTCTACCTTACCCGTTCCGTACCCCACTCTCAAGAGCTCTCTG 720  
DB 661 ATTGAAGACTTGAAGTCTCTACCTTACCCGTTCCGTACCCCACTCTCAAGAGCTCTCTG 720  
QY 721 AATATTTCTGTGTTTACCCCGAATATTTTACAGGTATCTGCAAGAGTCTCTTGGCCAG 780  
DB 721 AATATTTCTGTGTTTACCCCGAATATTTTACAGGTATCTGCAAGAGTCTCTTGGCCAG 780  
QY 781 GAGGAAAGCCAAAGTATCTGACCTTCAAGCAGATCCAGTTTGAAGGCCAATTTCAAG 840  
DB 781 GAGGAAAGCCAAAGTATCTGACCTTCAAGCAGATCCAGTTTGAAGGCCAATTTCAAG 840  
QY 841 GCAGTTCAACTTACTACGAATGATGCAATAAAACCACTGCTGTGTGAATTTGCAATT 900  
DB 841 GCAGTTCAACTTACTACGAATGATGCAATAAAACCACTGCTGTGTGAATTTGCAATT 900

QY 901 TCAAATACAGACTTTTCTTATCAGCTGAGATGCTTTCAGCGTATCTGCCCTAACAT 960  
DB 901 TCAAATACAGACTTTTCTTATCAGCTGAGATGCTTTCAGCGTATCTGCCCTAACAT 960  
QY 961 GATTTCAGGATCAAGAGCTTCTCAAGATTCAGATTCAGATTCAGATTCAGATTCAG 1020  
DB 961 GATTTCAGGATCAAGAGCTTCTCAAGATTCAGATTCAGATTCAGATTCAGATTCAG 1020  
QY 1021 GTTCCTTTGAAATTAAGGACAGACAAAGAAAGAGCTTACCTTACCCAGCATATA 1080  
DB 1021 GTTCCTTTGAAATTAAGGACAGACAAAGAAAGAGCTTACCTTACCCAGCATATA 1080  
QY 1081 CCGCGGATGTCTCTCCAGATTCATTTTACTGTGTCTTGAATTCGAGCAATTCCT 1140  
DB 1081 CCGCGGATGTCTCTCCAGATTCATTTTACTGTGTCTTGAATTCGAGCAATTCCT 1140  
QY 1141 AAAAAGCAATTTTGGAGCCCTTGTGGAATATACAGATGACAGTCTGAAAGCCAGG 1200  
DB 1141 AAAAAGCAATTTTGGAGCCCTTGTGGAATATACAGATGACAGTCTGAAAGCCAGG 1200  
QY 1201 CTACAGAGCTGTGAGTAAACAAAGGAGGACCGATTAATAGCCGTTTGTACAGATGCC 1260  
DB 1201 CTACAGAGCTGTGAGTAAACAAAGGAGGACCGATTAATAGCCGTTTGTACAGATGCC 1260  
QY 1261 TGTGCTGTGTGTGATCTCTCTCTGCTTCCCTTCCCTTCCGACGACCACTCAATCTC 1320  
DB 1261 TGTGCTGTGTGTGATCTCTCTCTGCTTCCCTTCCCTTCCGACGACCACTCAATCTC 1320  
QY 1321 CTGCTGAAATCTTCTTAACTTCAACCCAGACATATTCGTGTGCAAGCTCAAGTTTA 1380  
DB 1321 CTGCTGAAATCTTCTTAACTTCAACCCAGACATATTCGTGTGCAAGCTCAAGTTTA 1380  
QY 1381 TTTCAACCAAGAAAGCTCCATTTTGTCTTCAATTTGTGGAATTTGTCTACTGCCACA 1440  
DB 1381 TTTCAACCAAGAAAGCTCCATTTTGTCTTCAATTTGTGGAATTTGTCTACTGCCACA 1440  
QY 1441 ACAGAGTTCTGTGGAAGGAGATATGATACAGCTGTGCTGCTTGTGTGCTTCAATT 1500  
DB 1441 ACAGAGTTCTGTGGAAGGAGATATGATACAGCTGTGCTGCTTGTGTGCTTCAATT 1500  
QY 1501 CTTCAAGCAAACTATCATGATGATCCCATGAAAGACGCGGAAACCTTGCTCTTAAGATA 1560  
DB 1501 CTTCAAGCAAACTATCATGATGATCCCATGAAAGACGCGGAAACCTTGCTCTTAAGATA 1560  
QY 1561 TCCATCTTCTTGAACCAACAAATTTCTTCACTTACAGATGACCTTCAATCCCATC 1620  
DB 1561 TCCATCTTCTTGAACCAACAAATTTCTTCACTTACAGATGACCTTCAATCCCATC 1620  
QY 1621 AATAATGTGTGTCAAGAAACCGGATAGCCCGTTTATTTGGGTTCTTCAACATPAGAG 1680  
DB 1621 AATAATGTGTGTCAAGAAACCGGATAGCCCGTTTATTTGGGTTCTTCAACATPAGAG 1680  
QY 1681 AAATCCCAAGAACAAACCAAGATGAAATTTTGAAGCAATGTG---GTTTTTGGGTGC 1740  
DB 1681 AAATCCCAAGAACAAACCAAGATGAAATTTTGAAGCAATGTG---GTTTTTGGGTGC 1740  
QY 1741 AGCATTAAGATAGGATTAATCTATTTCAAGAAAGAGTCAAGATTTCTTAAAGCATGG 1800  
DB 1741 AGCATTAAGATAGGATTAATCTATTTCAAGAAAGAGTCAAGATTTCTTAAAGCATGG 1800  
QY 1738 AGCATTAAGATAGGATTAATCTATTTCAAGAAAGAGTCAAGATTTCTTAAAGCATGG 1797  
DB 1738 AGCATTAAGATAGGATTAATCTATTTCAAGAAAGAGTCAAGATTTCTTAAAGCATGG 1797  
QY 1801 ATCTTAATCTATTAAGGTTTCTTCTCAAGAGATGCTCTGTGGGAGGAGAAAGCC 1860  
DB 1801 ATCTTAATCTATTAAGGTTTCTTCTCAAGAGATGCTCTGTGGGAGGAGAAAGCC 1860  
QY 1798 ATCTTAATCTATTAAGGTTTCTTCTCAAGAGATGCTCTGTGGGAGGAGAAAGCC 1857  
DB 1798 ATCTTAATCTATTAAGGTTTCTTCTCAAGAGATGCTCTGTGGGAGGAGAAAGCC 1857  
QY 1861 CCAGCAAGATATATACAGACATCAAGCTTCAATGSCCAGAGGTCGCAATCTC 1920  
DB 1861 CCAGCAAGATATATACAGACATCAAGCTTCAATGSCCAGAGGTCGCAATCTC 1920  
QY 1858 CCAGCAAGATATATACAGACATCAAGCTTCAATGSCCAGAGGTCGCAATCTC 1917  
DB 1858 CCAGCAAGATATATACAGACATCAAGCTTCAATGSCCAGAGGTCGCAATCTC 1917  
QY 1921 CTCAGAGAAACGCAATATTTATGTGTGTGAGATGCAAGAAATATGSCCAAGATGTA 1980  
DB 1921 CTCAGAGAAACGCAATATTTATGTGTGTGAGATGCAAGAAATATGSCCAAGATGTA 1980  
QY 1918 CTCAGAGAAACGCAATATTTATGTGTGTGAGATGCAAGAAATATGSCCAAGATGTA 1977  
DB 1918 CTCAGAGAAACGCAATATTTATGTGTGTGAGATGCAAGAAATATGSCCAAGATGTA 1977  
QY 1981 CATGATGCCCTTGTGCAATTAATPAGCAAGAGGTTGAGTTGAAAACTPAGAAAGCATG 2040

|||||  
DB 1978 CATGATGCCCTGTGCAATAATATAGCAAGAGGTGAGTTGAAAACTAGAGCAATG 2037  
QY 2041 AAAACCCCTGCGCACTTTAAAGAGAAACCTACTCTCAGATTTTGTGCA 2094  
DB 2038 AAAACCTGCGCACTTTAAAGAGAAACCTACTCTCAGATTTTGTGCA 2091  
RESULT 10  
ADM43214  
ID ADM43214 standard; cDNA, 2091 BP.  
XX ADM43214;  
AC  
XX  
XX  
DT 03-JUN-2004 (first entry)  
XX  
XX Human methionine synthase reductase CDS del 1675-1678 variant.  
XX  
XX Human; ss; Methionine synthase reductase polypeptide; HsMTRR; cancer;  
KW cardiovascular disease; neural tube defect; hyperhomocysteinemia;  
XX chromosome 5p15.2-p15.3; SNP; single nucleotide polymorphism.  
XX  
OS Homo sapiens.  
FH  
FH Key Location/Qualifiers  
FT CDS 1..2091  
FT /tag= a  
FT /product= "HsMTRRdelR559"  
FT /partial  
FT /note= "No stop codon shown"  
FT /replace(66,A)  
FT variation /tag= b  
FT /standard name= "Single\_nucleotide\_polymorphism"  
FT /replace(110,A)  
FT /tag= c  
FT /standard name= "Single\_nucleotide\_polymorphism"  
FT variation /tag= d  
FT /replace(1675,AGAG)  
FT  
FT US2003082676-A1.  
XX  
XX  
XX  
XX PD 01-MAY-2003.  
XX  
XX PF 10-AUG-1999; 99US-00371347.  
XX  
XX PR 16-JAN-1998; 98US-0071622P.  
XX PR 15-JAN-1999; 99US-00232028.  
XX  
XX PA (GRAV/) GRAVEL R A.  
XX PA (ROZE/) ROZEN R.  
XX PA (LECL/) LECLERC D.  
XX PA (WILS/) WILSON A.  
XX PA (ROSE/) ROSENBLATT D.  
XX  
XX PI Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;  
XX  
XX WPI: 2003-576610/54.  
XX P-PSDB; ADM43215.  
XX  
XX PT New substantially pure nucleic acid encoding a mammalian methionine  
XX PT synthase reductase polypeptide, useful for diagnosis, preventing or  
XX PT treating conditions associated with altered methionine synthase activity,  
XX PT e.g. cancer.  
XX  
XX  
XX Discloure; SEQ ID NO 47; 26pp; English.  
XX  
XX  
XX The invention relates to a substantially pure nucleic acid that encodes a  
XX CC mammalian methionine synthase reductase polypeptide, HsMTRR, or that  
XX CC hybridizes at high stringency to a nucleic acid appearing as ADM43208 or  
XX CC ADM43209. Also included are a non-human animal where one or both genetic  
XX CC alleles encoding the methionine synthase reductase polypeptide are  
XX CC mutated, an antibody that specifically binds the above methionine  
XX CC synthase reductase polypeptide, a method of detecting the presence of the

CC above polypeptide, a method for detecting sequence variants for  
CC methionine synthase reductase in a mammal, methods of treating or  
CC preventing cancer (or cardiovascular disease or neural tube defects) in a  
CC subject, methods of screening for a compound that modulates methionine  
CC synthase reductase biological activity and a method for detecting an  
CC increased risk of developing a neural tube defect in a mammalian embryo  
CC or foetus. The nucleic acid is useful in diagnosing, preventing or  
CC treating conditions associated with altered methionine synthase activity,  
CC such as cancer, cardiovascular disease or neural tube defects, or in  
CC screening for a compound that modulates methionine synthase reductase  
CC biological activity. Naturally occurring variants of the polypeptide are  
CC also associated with hyperhomocysteinemia. The gene for HsMTRR is  
CC located on chromosome 5p15.2-p15.3. The present sequence is the coding  
CC sequence of a variant human HsMTRR cDNA.  
XX  
XX  
XX Sequence 2091 BP; 589 A; 489 C; 480 G; 533 T; 0 U; 0 Other;

Query Match 88.3%; Score 1851; DB 11; Length 2091;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2091; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 ATGAGAGGTTTCTGTACTATATGCTACACGACGAGGACGAGCAAGGCCATTCCGAGA 60  
DB 1 ATGAGAGGTTTCTGTACTATATGCTACACGACGAGGACGAGCAAGGCCATTCCGAGA 60  
QY 61 GAAATGTGTAGACAGCTGTGTACATGATTTTTCGAGATCTTCACTGATTAATGTA 120  
DB 61 GAAATGTGTAGACAGCTGTGTACATGATTTTTCGAGATCTTCACTGATTAATGTA 120  
QY 121 TCCGATTAAGTATGACCTTAACAAACGACTCTCTGTTGTTGTGTGTTTCTACACG 180  
DB 121 TCCGATTAAGTATGACCTTAACAAACGACTCTCTGTTGTTGTGTGTTTCTACACG 180  
QY 121 TCCGATTAAGTATGACCTTAACAAACGACTCTCTGTTGTTGTGTGTTTCTACACG 180  
DB 121 TCCGATTAAGTATGACCTTAACAAACGACTCTCTGTTGTTGTGTGTTTCTACACG 180  
QY 181 GGCACCGGAGACCCACCGACACAGCCGACAGTTTGTTAAGAAATACAAACCAACA 240  
DB 181 GGCACCGGAGACCCACCGACACAGCCGACAGTTTGTTAAGAAATACAAACCAACA 240  
QY 181 GGCACCGGAGACCCACCGACACAGCCGACAGTTTGTTAAGAAATACAAACCAACA 240  
DB 181 GGCACCGGAGACCCACCGACACAGCCGACAGTTTGTTAAGAAATACAAACCAACA 240  
QY 241 CTGCCGGTTGATTTCTGTTGCTCACTGCGGCTATGAGGTTCTCGGTTATTCAGA 300  
DB 241 CTGCCGGTTGATTTCTGTTGCTCACTGCGGCTATGAGGTTCTCGGTTATTCAGA 300  
QY 301 TACACCTACTTTTGGCAATGGGGGGAAGATATGATTAACGACTTCAAGAGCTTGAGCC 360  
DB 301 TACACCTACTTTTGGCAATGGGGGGAAGATATGATTAACGACTTCAAGAGCTTGAGCC 360  
QY 301 TACACCTACTTTTGGCAATGGGGGGAAGATATGATTAACGACTTCAAGAGCTTGAGCC 360  
DB 301 TACACCTACTTTTGGCAATGGGGGGAAGATATGATTAACGACTTCAAGAGCTTGAGCC 360  
QY 361 CCGCATTTCTATGACACTGACATGACATGACATGACATGACATGACATGACATGAC 420  
DB 361 CCGCATTTCTATGACACTGACATGACATGACATGACATGACATGACATGACATGAC 420  
QY 421 CCGTGAATGTCGACCTGCGGACGCTCAGAAAGCAATTTAAGTCAAGAGAGACAA 480  
DB 421 CCGTGAATGTCGACCTGCGGACGCTCAGAAAGCAATTTAAGTCAAGAGAGACAA 480  
QY 481 GAGAGATTAAGTGGGCACTCCCGGTGACATCCTGATCCTTGAGACAGACCTTGTG 540  
DB 481 GAGAGATTAAGTGGGCACTCCCGGTGACATCCTGATCCTTGAGACAGACCTTGTG 540  
QY 541 AAGTCAAGCTGCTACATTAATCTCAAGTCAAGCTTCTGAGATTCGATATTCAGGA 600  
DB 541 AAGTCAAGCTGCTACATTAATCTCAAGTCAAGCTTCTGAGATTCGATATTCAGGA 600  
QY 541 AAGTCAAGCTGCTACATTAATCTCAAGTCAAGCTTCTGAGATTCGATATTCAGGA 600  
DB 541 AAGTCAAGCTGCTACATTAATCTCAAGTCAAGCTTCTGAGATTCGATATTCAGGA 600  
QY 601 AAGAAAGATTCGAGGTTTGAAGCAAAATGACAGGAACAGAACCAATTCAGATGTGA 660  
DB 601 AAGAAAGATTCGAGGTTTGAAGCAAAATGACAGGAACAGAACCAATTCAGATGTGA 660  
QY 661 ATTGAAGACTTGAAGTCTCACTTACCCGTTGATACCTTCAAGACCTTCTGTG 720  
DB 661 ATTGAAGACTTGAAGTCTCACTTACCCGTTGATACCTTCAAGACCTTCTGTG 720  
QY 721 AATATTCCTGTTTACCCCAATATTTACAGTACATTCAGAGAGTCTTGTGCGCAG 780  
DB 721 AATATTCCTGTTTACCCCAATATTTACAGTACATTCAGAGAGTCTTGTGCGCAG 780



```
Qy 781 GAGAAAGCCAGATATCTGTACTTCAGCAGATCCAGTCTTTTCAAGTCCAAATTTCAAG 840
Db 781 GAGAAAGCCAGATATCTGTACTTCAGCAGATCCAGTCTTTTCAAGTCCAAATTTCAAG 840
Qy 841 GCGATTCACATTTACTAGAAATGATGCGATTAAGAACCACTGCGTGGAGAAATTTGACAT 900
Db 841 GCGATTCACATTTACTAGAAATGATGCGATTAAGAACCACTGCGTGGAGAAATTTGACAT 900
Qy 901 TCAATACAGACTTTTCTATTCAGCCTGAGATGCTTTCAGCGTATCTGCTTAAAGT 960
Db 901 TCAATACAGACTTTTCTATTCAGCCTGAGATGCTTTCAGCGTATCTGCTTAAAGT 960
Qy 961 GATTTGAGGTACAAAGCCTTACCTCAAGAGCTGACGCTTGAAGTAAAGAGCAGTGC 1020
Db 961 GATTTGAGGTACAAAGCCTTACCTCAAGAGCTGACGCTTGAAGTAAAGAGCAGTGC 1020
Qy 1021 GTGCTTTGAAATTAAGGAGACACAAAGAAAGAGACCTTACCCAGCATATA 1080
Db 1021 GTGCTTTGAAATTAAGGAGACACAAAGAAAGAGACCTTACCCAGCATATA 1080
Qy 1081 CTTGCGGAGATGTTCTCTCAGATTCATTTTACCTGCTGCTTGAATCCGAGCAATTCCT 1140
Db 1081 CTTGCGGAGATGTTCTCTCAGATTCATTTTACCTGCTGCTTGAATCCGAGCAATTCCT 1140
Qy 1141 AAAAAGCATTTTTCGAGCCCTTGTGACTATACAGTGAAGTCTGAAAAGCCGAGG 1200
Db 1141 AAAAAGCATTTTTCGAGCCCTTGTGACTATACAGTGAAGTCTGAAAAGCCGAGG 1200
Qy 1201 CTACAGAGCTGTGCAATTAACAGAGGAGCCGATTAATAGCCGCTTGTACGAGATGCC 1260
Db 1201 CTACAGAGCTGTGCAATTAACAGAGGAGCCGATTAATAGCCGCTTGTACGAGATGCC 1260
Qy 1261 TGTGCTGCTTGTGATCTCTCTCTGCTTCCCTTCTTCCAGCCAGCACTCAGTCTC 1320
Db 1261 TGTGCTGCTTGTGATCTCTCTCTGCTTCCCTTCTTCCAGCCAGCACTCAGTCTC 1320
Qy 1321 CTGCTCGAATCTTCTCTTAACTTCAACCCAGACATATTCGTGTGCAAGTCAAGTTTA 1380
Db 1321 CTGCTCGAATCTTCTCTTAACTTCAACCCAGACATATTCGTGTGCAAGTCAAGTTTA 1380
Qy 1381 TTTGACCCAGGAAAGCTCCATTTTGTCTTCAACATTTGAGAAATTTCTGTCTACTGCCACA 1440
Db 1381 TTTGACCCAGGAAAGCTCCATTTTGTCTTCAACATTTGAGAAATTTCTGTCTACTGCCACA 1440
Qy 1441 ACAGAGGTTCTGCGGAAAGGAGATATGTAACAGCTGCGCTTGTGTGCTTCAAGT 1500
Db 1441 ACAGAGGTTCTGCGGAAAGGAGATATGTAACAGCTGCGCTTGTGTGCTTCAAGT 1500
Qy 1501 CTTGAGCCAAACATATCATGCTCCCATGAGAGACGCGGAAAGCCCTGCGCTCTTAAAGATA 1560
Db 1501 CTTGAGCCAAACATATCATGCTCCCATGAGAGACGCGGAAAGCCCTGCGCTCTTAAAGATA 1560
Qy 1561 TCCATCTCTCTCGAACAACAAATTTCTTCCACTTACAGATGACCCCTCAATCCCATC 1620
Db 1561 TCCATCTCTCTCGAACAACAAATTTCTTCCACTTACAGATGACCCCTCAATCCCATC 1620
Qy 1621 ATTAATGCTGCTGCGAAGAACCGGATAGCCCGCTTATTTGCGTCTTCAACATAGAGAG 1680
Db 1621 ATTAATGCTGCTGCGAAGAACCGGATAGCCCGCTTATTTGCGTCTTCAACATAGAGAG 1680
Qy 1681 AAATCCCAAGAACACCCAGATGAGAAATTTTGGAGCATATGCTGTTTGGCTGC 1740
Db 1681 AAATCCCAAGAACACCCAGATGAGAAATTTTGGAGCATATGCTGTTTGGCTGC 1740
Qy 1741 AGGATTAAGATGAGGATTTATCTATTCAGAAAGAGCTCAGACATTTCTTAAAGATGGG 1800
Db 1741 AGGATTAAGATGAGGATTTATCTATTCAGAAAGAGCTCAGACATTTCTTAAAGATGGG 1800
Qy 1801 ATCTTAATCTATTAAGGATTTCTTCTCAAGAGATGCTCTGTTGGGAGAGAGAGCC 1860
Db 1801 ATCTTAATCTATTAAGGATTTCTTCTCAAGAGATGCTCTGTTGGGAGAGAGAGCC 1860
Qy 1861 CAGCAAGATATGTAAGAGACATTCAGCTTACGCTTACGCGCAGAGTGGCAGATATCTTC 1920
Db 1861 CAGCAAGATATGTAAGAGACATTCAGCTTACGCTTACGCGCAGAGTGGCAGATATCTTC 1920
```

```
Db 1858 CCAGCAAGATATGTAAGAGACATTCAGCTTACGCTTACGCGCAGAGTGGCAGATATCTTC 1917
Qy 1921 CTCAGAGAGAGGCGCATATTTATGTGTGAGATGCAAGAAATATGCGCAAGATGTA 1980
Db 1918 CTCAGAGAGAGGCGCATATTTATGTGTGAGATGCAAGAAATATGCGCAAGATGTA 1977
Qy 1981 CATGATGCCCTTGTGCAATTAATAGCAAGAGGTTGAGTTGAAAACTAGAAAGCAATG 2040
Db 1978 CATGATGCCCTTGTGCAATTAATAGCAAGAGGTTGAGTTGAAAACTAGAAAGCAATG 2037
Qy 2041 AAAACCTGCGCACTTAAAGAGAAAGAGCTACCTCAGGATATTTGTCA 2094
Db 2038 AAAACCTGCGCACTTAAAGAGAAAGAGCTACCTCAGGATATTTGTCA 2091
```

## RESULT 11

AA58977 standard; DNA; 3256 BP.

AA58977;

07-NOV-2000 (first entry)

A human methionine synthase reductase DNA sequence with polymorphism.

Human; methionine synthase reductase; MTRR; cancer;

cardiovascular disease; Down's Syndrome; neural tube defect;

premature coronary artery disease; ss.

Homo sapiens.

WO200042196-A2.

20-JUL-2000.

14-JAN-2000; 2000MO-IB000209.

15-JAN-1999; 99US-00232028.

10-AUG-1999; 99US-00371347.

(UYNM-) UNIV MCGILL.

Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;

MPI; 2000-46131/40.

Claim 8; Page; 85pp; English.

The present sequence represents a human methionine synthase reductase

(MTRR) DNA sequence, with a polymorphism comprising of a deletion of

nucleotides 1726-1728. Inhibitors of MTRR polypeptide and polynucleotide

are used for treating or preventing cancer, cardiovascular disease,

Down's Syndrome or neural tube defects in a subject. The cardiovascular

disease is premature coronary artery disease. The compounds are detected

by methods which screen for modulators of MTRR biological activity. MTRR

polypeptide or nucleic acid is examined for the presence of a

polymorphism in the parents or the embryo or foetus, and the information

used for detecting an increased risk of an embryo or foetus developing

cancer, cardiovascular disease, Down's Syndrome or neural tube defects.

note: the present sequence does not appear in the specification; it was

created using information provided

Sequence 3256 BP; 943 A; 705 C; 662 G; 946 T; 0 U; 0 Other;

Query Match 83.5%; Score 1752; DB 3; Length 3256;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 2092; Conservative 0; Mismatches 2; Indels 3; Gaps 1;



QY 1 ATGAGAGGTTCTGTACTATATGCTACACAGCAGGAGCAGGCAAAAGCCATCGAGAA 60  
DB 80 ATGAGAGGTTCTGTACTATATGCTACACAGCAGGAGCAGGCAAAAGCCATCGAGAA 139  
QY 61 GAAATGTGAGCAGCTGTGTACATGATTTTCTGAGATCTTCACTGTATTAGTAA 120  
DB 140 GAAATGTGAGCAGCTGTGTACATGATTTTCTGAGATCTTCACTGTATTAGTAA 199  
QY 121 TCCGAAATGATGACCTTAAACCCGAAACAGCTCTCTGTTGTGTGTTCTACAG 180  
DB 200 TCCGAAATGATGACCTTAAACCCGAAACAGCTCTCTGTTGTGTGTTCTACAG 259  
QY 181 GGCACCGGAGACCCACCCGACACAGCCCGAGTTGTTAAGGAATACAGAACCA 240  
DB 260 GGCACCGGAGACCCACCCGACACAGCCCGAGTTGTTAAGGAATACAGAACCA 319  
QY 241 CTGCGGTTGATTTCTTGTCTCACTGCGGATGAGTTACTGGGTCCTGATTCAGAA 300  
DB 320 CTGCGGTTGATTTCTTGTCTCACTGCGGATGAGTTACTGGGTCCTGATTCAGAA 379  
QY 301 TACACTACTTTTGTGCAATGGGGGAGATTAATGATTAACGACTTCAAGAGCTTGAGCC 360  
DB 380 TACACTACTTTTGTGCAATGGGGGAGATTAATGATTAACGACTTCAAGAGCTTGAGCC 439  
QY 361 CGCATTTTCTATGACATGACATGACATGACATGACATGACATGACATGACATGAC 420  
DB 440 CGCATTTTCTATGACATGACATGACATGACATGACATGACATGACATGACATGAC 499  
QY 421 CCGTGATTTCTGTGACCTGTGACAGCCCTTCAAGAACATTTTGTGACAGAGCA 480  
DB 500 CCGTGATTTCTGTGACCTGTGACAGCCCTTCAAGAACATTTTGTGACAGAGCA 559  
QY 481 GAGGAAATAGTGGCGCATCTCCGTTGGCATCTCTGCACTCTTGAGAGCAGCTTGTG 540  
DB 560 GAGGAAATAGTGGCGCATCTCCGTTGGCATCTCTGCACTCTTGAGAGCAGCTTGTG 619  
QY 541 AAGTCAGAGCTGTACATGATGATCTCAAGTCGAGCTTGTGATTCAGATTCAGAG 600  
DB 620 AAGTCAGAGCTGTACATGATGATCTCAAGTCGAGCTTGTGATTCAGATTCAGAG 679  
QY 601 AGAAGATTTCTGAGGTTTGAAGCAAAATGACAGTGAACAGCAACCAATTCATGTTGTA 660  
DB 680 AGAAGATTTCTGAGGTTTGAAGCAAAATGACAGTGAACAGCAACCAATTCATGTTGTA 729  
QY 661 ATTGAAGCTTTGAGTCTCTCACTTACCCGTTGGTACCCCACTCTCAAGCTCTCTG 720  
DB 740 ATTGAAGCTTTGAGTCTCTCACTTACCCGTTGGTACCCCACTCTCTCAAGCTCTCTG 799  
QY 721 AATATTCGTGTTTACCCCGAATTTTACAGATCATCTGACAGAGTCTTGTGGCAG 780  
DB 800 AATATTCGTGTTTACCCCGAATTTTACAGATCATCTGACAGAGTCTTGTGGCAG 859  
QY 781 GAGGAAAGCCAGATCTGTGACCTTCAAGATTCAGTTTTCAGTGCATTTCAAG 840  
DB 860 GAGGAAAGCCAGATCTGTGACCTTCAAGATTCAGTTTTCAGTGCATTTCAAG 919  
QY 841 GCAGTTCATTTACATGAGTATGATGCTATTAACCACTCTGCTGTGATTTGACAT 900  
DB 920 GCAGTTCATTTACATGAGTATGATGCTATTAACCACTCTGCTGTGATTTGACAT 979  
QY 901 TCAATATGACATTTTCTTCTATGAGCTGAGATGCTTCAAGGATTCGCTTAAAGT 960  
DB 980 TCAATATGACATTTTCTTCTATGAGCTGAGATGCTTCAAGGATTCGCTTAAAGT 1039  
QY 961 GATTTGAGGTAACAAAGCTTCTCAAGATCTGACAGCTTGAAGATTAAGAGACCTGC 1020  
DB 1040 GATTTGAGGTAACAAAGCTTCTCAAGATCTGACAGCTTGAAGATTAAGAGACCTGC 1099  
QY 1021 GTCTTTTGAATAAAGGAGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080  
DB 1100 GTCTTTTGAATAAAGGAGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1159  
QY 1081 CTGCGGAGATTTCTCTCAAGTTCATTTTACCTGATGTTGAAATTCGAGCAATTCCT 1140

DB 1160 CTGCGGAGATTTCTCTCAAGTTCATTTTACCTGATGTTGAAATTCGAGCAATTCCT 1219  
QY 1141 AAAAAGCATTTTGTGAGACCTTGTGACATTAACAGAGACAGCTTAAAGCCAG 1200  
DB 1220 AAAAAGCATTTTGTGAGACCTTGTGACATTAACAGAGACAGCTTAAAGCCAG 1279  
QY 1201 CTACAGAGCTGTGACATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260  
DB 1280 CTACAGAGCTGTGACATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1339  
QY 1261 TGTGCTGTGTGTGATTCCTCTCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1320  
DB 1340 TGTGCTGTGTGTGATTCCT 1399  
QY 1321 CTGCTGAAACATTTCTTAACTTCAACCCAGACATATTCGTGTGACAGCTTAA 1380  
DB 1400 CTGCTGAAACATTTCTTAACTTCAACCCAGACATATTCGTGTGACAGCTTAA 1459  
QY 1381 TTTTACCCGAGAAAGCTCAATTTGTCTTCAATTTGAGAAATTTCTCTACTGACACA 1440  
DB 1460 TTTTACCCGAGAAAGCTCAATTTGTCTTCAATTTGAGAAATTTCTCTACTGACACA 1519  
QY 1441 ACAGAGTTCTGCGAAGGAGATGATACAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
DB 1520 ACAGAGTTCTGCGAAGGAGATGATACAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1579  
QY 1501 CTTCAGCCAAACATACATGATCTCCATGAAAGACAGGAGAGAGAGAGAGAGAGAG 1560  
DB 1580 CTTCAGCCAAACATACATGATCTCCATGAAAGACAGGAGAGAGAGAGAGAGAGAG 1639  
QY 1561 TCCATCTCTCTGAAACAAATTTCTTCACTTACAGATGACAGCTTCAATCCCATC 1620  
DB 1640 TCCATCTCTCTGAAACAAATTTCTTCACTTACAGATGACAGCTTCAATCCCATC 1699  
QY 1621 ATATGTTGTTCTGAGAACAGCCGATAGCCGTTTATTTGTTTCTTCAACATAGAGAG 1680  
DB 1700 ATATGTTGTTCTGAGAACAGCCGATAGCCGTTTATTTGTTTCTTCAACATAGAGAG 1756  
QY 1681 AAATCTCAAGAACACACCCAGATGGAATTTTGAAGCATGTTGTTTGTGCTGC 1740  
DB 1757 AAATCTCAAGAACACACCCAGATGGAATTTTGAAGCATGTTGTTTGTGCTGC 1816  
QY 1741 AGGCAATAGATGAGATTAATTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800  
DB 1817 AGGCAATAGATGAGATTAATTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1876  
QY 1801 ATCTTAATCATTAAGGTTTCTTCTCAAGAGATGCTCTGTTGGAGAGAGAGAGAG 1860  
DB 1877 ATCTTAATCATTAAGGTTTCTTCTCAAGAGATGCTCTGTTGGAGAGAGAGAGAG 1936  
QY 1861 CCAGCAAGATGATACAGACACATCCAGCTTCAATGCGACAGAGTGGCGAGATCTC 1920  
DB 1937 CCAGCAAGATGATACAGACACATCCAGCTTCAATGCGACAGAGTGGCGAGATCTC 1996  
QY 1921 CTCCAGGAAAGGCGCATTTTATGTTGTGAGAGATGCAAGAAATATGCGCAAGATGTA 1980  
DB 1997 CTCCAGGAAAGGCGCATTTTATGTTGTGAGAGATGCAAGAAATATGCGCAAGATGTA 2056  
QY 1981 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGTGTGAGTTGAAAAATAGAGCAATG 2040  
DB 2057 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGTGTGAGTTGAAAAATAGAGCAATG 2116  
QY 2041 AAAACCTTGCCCATTTTAAAGAGAAAAAGCTTCAAGATTTTGTGCTATA 2097  
DB 2117 AAAACCTTGCCCATTTTAAAGAGAAAAAGCTTCAAGATTTTGTGCTATA 2173

RESULT 12  
AA58976  
ID AA58976 standard; DNA; 3255 BP.  
XX  
AC AA58976;

XX 07-NOV-2000 (first entry)  
XX A human methionine synthase reductase DNA sequence with polymorphism.  
DE Human; methionine synthase reductase; MTRR; cancer;  
KW cardiovascular disease; Down's Syndrome; neural tube defect;  
KM premature coronary artery disease; ss.  
XX Homo sapiens.  
XX MO200042196-A2.  
XX 20-JUL-2000.  
XX 14-JAN-2000; 2000MO-IB000209.  
XX 15-JAN-1999; 99US-00232028.  
XX 10-AUG-1999; 99US-00371347.  
XX (UYMC-) UNIV MCGILL.  
XX PA  
XX PI Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;  
XX MPI; 2000-466131/40.  
XX Mammalian methionine synthase reductase nucleic acid used for detecting  
XX an increased risk of developing a neural tube defect. Down's Syndrome or  
XX cardiovascular disease in a mammalian embryo or fetus.  
XX Claim 7; Page; 85pp; English.  
XX  
XX The present sequence represents a human methionine synthase reductase  
XX (MTRR) DNA sequence, with a polymorphism comprising of a deletion of  
XX nucleotides 1675-1678. Inhibitors of MTRR polypeptide and polynucleotide  
XX are used for treating or preventing cancer, cardiovascular disease,  
XX Down's Syndrome or neural tube defects in a subject. The cardiovascular  
XX disease is premature coronary artery disease. The compounds are detected  
XX by methods which screen for modulators of MTRR biological activity. MTRR  
XX polypeptide or nucleic acid is examined for the presence of a  
XX polymorphism in the parents or the embryo or foetus, and the information  
XX used for detecting an increased risk of an embryo or foetus developing  
XX cancer, cardiovascular disease, Down's Syndrome or neural tube defects.  
XX note: the present sequence does not appear in the specification; it was  
XX created using information provided  
XX  
XX Sequence 3255 BP; 942 A; 704 C; 663 G; 946 T; 0 U; 0 Other;  
XX  
XX Query Match 80.6%; Score 1691; DB 3; Length 3255;  
XX Best Local Similarity 99.7%; Pred. No. 0;  
XX Matches 2091; Conservative 0; Mismatches 2; Indels 4; Gaps 1;  
XX  
QY 1 ATGAGAGAGTTCTGTTACTATATGCTACACAGAGGAGACGAAAGCCATGCGAGAA 60  
DB 80 ATGAGAGAGTTCTGTTACTATATGCTACACAGAGGAGACGAAAGCCATGCGAGAA 139  
QY 61 GAAATGTGTGACAGCTGTGTGATCATGATTTTCTGCAGATCTTCACTGATTAAGTAA 120  
DB 140 GAAATGTGTGACAGCTGTGTGATCATGATTTTCTGCAGATCTTCACTGATTAAGTAA 199  
QY 121 TCCGATTAAGTATGACCTTAAAAACCGAAACAGCTCCTCTGTTGTTGTGTTTCTAACACG 180  
DB 200 TCCGATTAAGTATGACCTTAAAAACCGAAACAGCTCCTCTGTTGTTGTGTTTCTAACACG 259  
QY 181 GGCACCGGAGACCAACCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACA 240  
DB 260 GGCACCGGAGACCAACCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACA 319  
QY 241 CTGCGGATTTGATTTCTTTGCTCACTGCGGTATGGTTTACTGGCTCTCGGATTTCAAGAA 300  
DB 320 CTGCGGATTTGATTTCTTTGCTCACTGCGGTATGGTTTACTGGCTCTCGGATTTCAAGAA 379  
QY 301 TACACCTTACTTTGCAATGGGGGAAAGATTAATTGATTAACGATTCAGAGCTTGAGCC 360

DB 380 TACACCTTACTTTGCAATGGGGGAAAGATTAATTGATTAACGATTCAGAGCTTGAGCC 439  
QY 361 CGGATTTCTATGACACTGAGACATGACAGATGCTGTAGTTAGACTTGTGTTGAG 420  
DB 440 CGGATTTCTATGACACTGAGACATGACAGATGCTGTAGTTAGACTTGTGTTGAG 499  
QY 421 CCGTGATTTGCTGACTCTGCGCCAGCCCTCAGAAAGCATTTTATGTCAGACAGACAA 480  
DB 500 CCGTGATTTGCTGACTCTGCGCCAGCCCTCAGAAAGCATTTTATGTCAGACAGACAA 559  
QY 481 GAGGAGATAGTGGCGCACTCCCGGTGGATACCTCGATCTTGTAGGACAGACTTGTG 540  
DB 560 GAGGAGATAGTGGCGCACTCCCGGTGGATACCTCGATCTTGTAGGACAGACTTGTG 619  
QY 541 AAGTCAGACTGCTACACATTTGATATCTCAAGTCGACTTCTGAGATTGATTCAGGA 600  
DB 620 AAGTCAGACTGCTACACATTTGATATCTCAAGTCGACTTCTGAGATTGATTCAGGA 679  
QY 601 AGAAGGATTTGAGGTTTGAAGCAAAATGCAGTGAACAGCAACCAATCCATGTTGTA 660  
DB 680 AGAAGGATTTGAGGTTTGAAGCAAAATGCAGTGAACAGCAACCAATCCATGTTGTA 739  
QY 661 ATTGAAGACTTGAAGTCCCTCACTTACCCTGTGGTACCCCACTCCACAGGCTCTG 720  
DB 740 ATTGAAGACTTGAAGTCCCTCACTTACCCTGTGGTACCCCACTCCACAGGCTCTG 799  
QY 721 AATATTCCTGTTTATCCCGCAGAAATTTTACAGGTATCATCTGACAGAGTCTTGGCCAG 780  
DB 800 AATATTCCTGTTTATCCCGCAGAAATTTTACAGGTATCATCTGACAGAGTCTTGGCCAG 859  
QY 781 GAGGAAAGCCAGATATCTGTGACTTACGACGATCCAGTTTTCAGTCCCATTTCAAG 840  
DB 860 GAGGAAAGCCAGATATCTGTGACTTACGACGATCCAGTTTTCAGTCCCATTTCAAG 919  
QY 841 GCGATTCATCTTACATGCAATGATGCAATTAACCACTGTGTGATTAAGTGGACAT 900  
DB 920 GCGATTCATCTTACATGCAATGATGCAATTAACCACTGTGTGATTAAGTGGACAT 979  
QY 901 TCAATATCAGACTTTCTCTATCAGCTGTGAGATGCTTACGCTGTATCTGCCCTTAACGT 960  
DB 980 TCAATATCAGACTTTCTCTATCAGCTGTGAGATGCTTACGCTGTATCTGCCCTTAACGT 1039  
QY 961 GATTCGAGGTACAAAGCCTTCTCAAGACTGCAAGCTGTGAATTAAGAGAGCACTGC 1020  
DB 1040 GATTCGAGGTACAAAGCCTTCTCAAGACTGCAAGCTGTGAATTAAGAGAGCACTGC 1099  
QY 1021 GTCCCTTTGAAATTAAGGACACAGAAAGAGAGCTTACCTTACCCGACATATA 1080  
DB 1100 GTCCCTTTGAAATTAAGGACACAGAAAGAGAGCTTACCTTACCCGACATATA 1159  
QY 1081 CTTGCGGAGATGTTCTCTCCAGTTCAATTTTAACTGTGTGTAATCCGAGCAATTCCT 1140  
DB 1160 CTTGCGGAGATGTTCTCTCCAGTTCAATTTTAACTGTGTGTAATCCGAGCAATTCCT 1219  
QY 1141 AAAAAGCATTTTGGAGAGCCCTTGTGACTATACAGTACAGTGTGTAAGCCGAGG 1200  
DB 1220 AAAAAGCATTTTGGAGAGCCCTTGTGACTATACAGTACAGTGTGTAAGCCGAGG 1279  
QY 1201 CTACAGAGCTGTGCAATTAACAAAGGGGACCGGATTAATAGCCGCTTGTAGAGAGTGC 1260  
DB 1280 CTACAGAGCTGTGCAATTAACAAAGGGGACCGGATTAATAGCCGCTTGTAGAGAGTGC 1339  
QY 1261 TGTGCTGCTGTTGTGATCTCTCTCGCTTCCCTTCTTCCAGCAGCAGCACTCAGTCTC 1320  
DB 1340 TGTGCTGCTGTTGTGATCTCTCTCGCTTCCCTTCTTCCAGCAGCAGCACTCAGTCTC 1399  
QY 1321 CTCTCGAATCATTTCTTAACTTCAACCCAGACCATATTTGTGTGCAAGCTCAAGTTTA 1380  
DB 1400 CTCTCGAATCATTTCTTAACTTCAACCCAGACCATATTTGTGTGCAAGCTCAAGTTTA 1459  
QY 1381 TTTCACCCAGAGAGCTCCATTTTGTCTTCAAACTGTGGAATTTCTGTCTACCTGACCA 1440

```

Db 1460 TTTCAACCCAGAAAGCTCCATTGTTCTTCAACATTGTGAAATTTCTGTCTACTGCCACA 1519
Qy 1441 ACAGAGTTCTGGAGGAGGATATGATAGAGCTGGCTGGCTGTTGGTCTTCAGTT 1500
Db 1520 ACAGAGTTCTGGAGGAGGATATGATAGAGCTGGCTGGCTGTTGGTCTTCAGTT 1579
Qy 1501 CTTGAGCCAAATATCATCATCCCATGAGAGCAGCGGAAAGCCCTGGCTCTAAGATA 1560
Db 1580 CTTGAGCCAAATATCATCATCCCATGAGAGCAGCGGAAAGCCCTGGCTCTAAGATA 1639
Qy 1561 TCCATCTCTCTCGAACAACAATTTCTCACTTACCAAGATGACCCCTCAATCCCATC 1620
Db 1640 TCCATCTCTCTCGAACAACAATTTCTCACTTACCAAGATGACCCCTCAATCCCATC 1695
Qy 1621 ATAAATGTTGGTCCAGAAACCGGCAATGCGCTTATGTTGGTCTTACAACATGAGAG 1680
Db 1696 ATAAATGTTGGTCCAGAAACCGGCAATGCGCTTATGTTGGTCTTACAACATGAGAG 1755
Qy 1681 AAATCTCAGAAACAACCCAGATGGAATTTTGGAGCAATGTGTTGTTTGGCTGC 1740
Db 1756 AAATCTCAGAAACAACCCAGATGGAATTTTGGAGCAATGTGTTGTTTGGCTGC 1815
Qy 1741 AGGCAATAGGATAGGATATCTATTCTAGAAAAAGCTCAACATTTCTTAAGCATGCG 1800
Db 1816 AGGCAATAGGATAGGATATCTATTCTAGAAAAAGCTCAACATTTCTTAAGCATGCG 1875
Qy 1801 ATCTTAACATCTAAAGTTTCTCTGAGAGATGCTCTGTTGGGAGAGAGAGAGCC 1860
Db 1876 ATCTTAACATCTAAAGTTTCTCTGAGAGATGCTCTGTTGGGAGAGAGAGAGCC 1935
Qy 1861 CCAGCAAAATATGATCAAGACAACATCCAGCTTCATGCGCAGCAGGTGCGAATCTCTC 1920
Db 1936 CCAGCAAAATATGATCAAGACAACATCCAGCTTCATGCGCAGCAGGTGCGAATCTCTC 1995
Qy 1921 CTCGAGAGAGACGGCCATATTTATGTGTGTGAGATGCAAGAAATATGCGCAAGATGTA 1980
Db 1996 CTCGAGAGAGACGGCCATATTTATGTGTGTGAGATGCAAGAAATATGCGCAAGATGTA 2055
Qy 1981 CATGATGCTCTGTGCAATATTAAGCAAGAGGTTGAGGTTGAAAAACTAGAGCAATG 2040
Db 2056 CATGATGCTCTGTGCAATATTAAGCAAGAGGTTGAGGTTGAAAAACTAGAGCAATG 2115
Qy 2041 AAAACCTGCGCACTTTAAAGAAAGAAAAAGCTACCTTCAGGATTTTGGTCATTA 2097
Db 2116 AAAACCTGCGCACTTTAAAGAAAGAAAAAGCTACCTTCAGGATTTTGGTCATTA 2172

```

## RESULT 13

ADQ39029  
ID ADQ39029 standard; DNA; 3256 BP.

ADQ39029;

18-NOV-2004 (first entry)

Human SNP containing myocardial infarction-associated gene, SEQ ID 692.

Myocardial infarction; detection; single nucleotide polymorphism; SNP;  
cardiac; gene therapy; human; gene; ds.

Homo sapiens.

WO2004058052-A2.

15-JUL-2004.

22-DEC-2003; 2003WO-US040978.

20-DEC-2002; 2002US-0434778P.

10-MAR-2003; 2003US-0453135P.

30-APR-2003; 2003US-0466412P.

23-SEP-2003; 2003US-0504955P.

PA (Appl.) APPLERA CORP.  
XX  
XX  
PI Carcilli M, Devlin JF, Iakoubova O;  
XX  
XX  
DR WPI; 2004-533949/51.  
P-PSDB; ADQ39857.  
XX  
XX  
PT Identifying an individual who has an altered risk for developing  
PT myocardial infarction by detecting a single nucleotide polymorphism in  
PT the individual's nucleic acids.  
XX  
XX  
PS Claim 7; SEQ ID NO 692; 145bp; English.

CC The invention relates to a novel method for identifying an individual who  
CC has an altered risk for developing myocardial infarction. The method  
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of  
CC the nucleotide sequences given in the specification in the individual's  
CC nucleic acids, where the presence of the SNP is correlated with an  
CC altered risk for myocardial infarction in the individual. The invention  
CC further comprises: an isolated nucleic acid molecule comprising at least  
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in  
CC the specification or its complement and encoding any one of the amino  
CC acid sequences given in the specification; an isolated polypeptide  
CC comprising an amino acid sequence given in the specification; an antibody  
CC that specifically binds to the polypeptide or its antigen-binding  
CC fragment; an amplified polynucleotide containing an SNP given in the  
CC specification and which is between about 16 and 1000 nucleotides in  
CC length; a kit for detecting an SNP in a nucleic acid, comprising the  
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a  
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a  
CC method for identifying an agent useful in treating or preventing  
CC myocardial infarction. The novel detection method has cardiac activity.  
CC The nucleic acids of the invention may be used in gene therapy. The  
CC method is useful in identifying an individual who has an increased or  
CC decreased risk for developing myocardial infarction and for preparing a  
CC composition for treating or preventing myocardial infarction. This  
CC polynucleotide sequence represents a human myocardial infarction-  
CC associated gene containing one or more SNP's of the invention. Note: This  
CC sequence was not shown in the specification. The sequence has come from  
CC an electronic sequence listing downloaded from the WIPO website.  
XX

Sequence 3256 BP; 927 A; 691 C; 669 G; 940 T; 0 U; 29 Other;

Query Match 50.6%; Score 1062; DB 13; Length 3256;  
Best Local Similarity 99.1%; Pred. No. 0;  
Matches 2012; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

```

Qy 67 TGTGAGCAAGCTGTGTATCATGATTTTCTGCAGATCTTCACTGATTAAGTAATCCGAT 126
Db 160 TGTGAGCAAGCTGTGTATCATGATTTTCTGCAGATCTTCACTGATTAAGTAATCCGAT 219
Qy 127 AAGTATGACCTAAACCGAAACAGCTCTCTTGTGTGTGTGTTCTACACGGGACCC 186
Db 220 AAGTATGACCTAAACCGAAACAGCTCTCTTGTGTGTGTGTTCTACACGGGACCC 279
Qy 187 GGAGACCCACCGGACACAGCCCGCAAGTTGTTAAGAAATACAGAACCAAACTGCGCG 246
Db 280 GGAGACCCACCGGACACAGCCCGCAAGTTGTTAAGAAATACAGAACCAAACTGCGCG 339
Qy 247 GTTGATTTCTTTGCTACCTGGGTATGGGTACTGGGTCTGGGTATTCAGAAATACACC 306
Db 340 GTTGATTTCTTTGCTACCTGGGTATGGGTACTGGGTCTGGGTATTCAGAAATACACC 399
Qy 307 TACTTTTGAATGGGGGGAAGATTAATGATTAACGATTCAGAGCTTGAGCCGGCAT 366
Db 400 TACTTTTGAATGGGGGGAAGATTAATGATTAACGATTCAGAGCTTGAGCCGGCAT 459
Qy 367 TTCTATGACATGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 426
Db 460 TTCTATGACATGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 519
Qy 427 ATTGCTGACTCTGGCCAGCCCTCAGAAAGATTTTAAAGTCAAGAGAGAGCAAGAGAG 486

```

Db 520 ATTGTCGACCTGCGCAGACCCCTCAGAAAGCATTTTAGTCAAGCAGAGGACAAGAGGAG 579  
Qy 487 ATAAAGGCGCATCTCCCGGTGCGATCAGTCATCTTGAGAGCAGACCTGTGAAGTCA 546  
Db 580 ATTAAGGCGCATCTCCCGGTGCGATCAGTCATCTTGAGAGGACAAGCCTGTGAAGTCA 639  
Qy 547 GAGCTGCTACACATTTGAATCTCAAGTCGAGCTTCTGAGATTGATATTCAGGAAGAAAG 606  
Db 640 GAGCTGCTACACATTTGAATCTCAAGTCGAGCTTCTGAGATTGATATTCAGGAAGAAAG 699  
Qy 607 GATTGAGGTTTGAAGCAAAATGAGAGAGCAAGCAACCAATCCAAATGTTGAATTTGA 666  
Db 700 GATTGAGGTTTGAAGCAAAATGAGAGAGCAAGCAACCAATCCAAATGTTGAATTTGA 759  
Qy 667 GACTTGAAGTCCTACATTAACCCGTGCGTACCCCACTCTCAGAGCTCTGAAATATT 726  
Db 760 GACTTGAAGTCCTACATTAACCCGTGCGTACCCCACTCTCAGAGCTCTGAAATATT 819  
Qy 727 CCGTGTATACCCCGAGAAATTTTACAGGTACATCTGAGAGAGTCTCTTGCGCAGAGGAA 786  
Db 820 CCGTGTATACCCCGAGAAATTTTACAGGTACATCTGAGAGAGTCTCTTGCGCAGAGGAA 879  
Qy 787 AGCCAAGTATCTGAGACTTGACAGAGATCCAGTTTTCAGTGGCAATTTCAAGGCGAGT 846  
Db 880 AGCCAAGTATCTGAGACTTGACAGAGATCCAGTTTTCAGTGGCAATTTCAAGGCGAGT 939  
Qy 847 CAATTAATCAAGATGATGCGCATTAATAAACCACTCTGCTGTAGAAATTTGACATTTCAAT 906  
Db 940 CAATTAATCAAGATGATGCGCATTAATAAACCACTCTGCTGTAGAAATTTGACATTTCAAT 999  
Qy 907 ACAAGCTTTTCTATCAGCTCGAGATGCTTACGCGTATCTGCGCTTAACAGTATTTCT 966  
Db 1000 ACAAGCTTTTCTATCAGCTCGAGATGCTTACGCGTATCTGCGCTTAACAGTATTTCT 1059  
Qy 967 GAGGTCAAAAGCCTACTCCAAAGCTGACAGTTGAAGATTAAGAGAGCACTGCGTCTT 1026  
Db 1060 GAGGTCAAAAGCCTACTCCAAAGCTGACAGTTGAAGATTAAGAGAGCACTGCGTCTT 1119  
Qy 1027 TTGAAATTAAGGACGACACAAAGAGAAAGAGCTACTTACCCGACATATACCTGCG 1086  
Db 1120 TTGAAATTAAGGACGACACAAAGAGAAAGAGCTACTTACCCGACATATACCTGCG 1179  
Qy 1087 GATGTTCTCTCAAGTCAATTTTAACTGCTGTCTTGAATTCGAGCAATTTCTAATAAG 1146  
Db 1180 GATGTTCTCTCAAGTCAATTTTAACTGCTGTCTTGAATTCGAGCAATTTCTAATAAG 1239  
Qy 1147 GCATTTTGGAGGCTTGTGGAATATACAGTGAAGTGTGTAAGAAAGCGAGGCTACAG 1206  
Db 1240 GCATTTTTCAGGCTTGTGGAATATACAGTGAAGTGTGTAAGAAAGCGAGGCTACAG 1299  
Qy 1207 GAGCTGTGAGTAACAAGAGGCGCAGATTATAGCCGCTTGTATAGAGATGCTGTGCC 1266  
Db 1300 GAGCTGTGAGTAACAAGAGGCGCAGATTATAGCTTGTATAGAGATGCTGTGCC 1359  
Qy 1267 TGCCTTGTGATCTCTCTCTGCTTTCCTTCTTGCCAGGCACTCAAGTCTCTGCTC 1326  
Db 1360 TGCCTTGTGATCTCTCTCTGCTTTCCTTCTTGCCAGGCACTCAAGTCTCTGCTC 1419  
Qy 1327 GAACATCTTCTTAACCTTCAACCCAGACATATTCGTTGAGAGCTCAAGTTTATTTAC 1386  
Db 1420 GAACATCTTCTTAACTTCAACCCAGACATATTCGTTGAGAGCTCAAGTTTATTTAC 1479  
Qy 1387 CAGAGAAAGCTCAATTTTGTCTTCAACATTTGTGAAATTTCTGTCTACCTGCCAACAAG 1446  
Db 1480 CAGAGAAAGCTCAATTTTGTCTTCAACATTTGTGAAATTTCTGTCTACCTGCCAACAAG 1539  
Qy 1447 GTTCTGCGAAGGAGATGTAAGGCTGCTGCTTGTGTTGCTTCACTTCTTACAG 1506  
Db 1540 GTTCTGCGAAGGAGATGTAAGGCTGCTGCTTGTGTTGCTTCACTTCTTACAG 1599  
Qy 1507 CCAAACTATCATTCATCCATGAAGACAGCGGAAAGCCCTGAGCTCCCTAAGATATCCATC 1566  
Db 1600 CCAAACTATCATTCATCCATGAAGACAGCGGAAAGCCCTGAGCTCCCTAAGATATCCATC 1659

Qy 1567 TCTCCTCGAACAACAATTTCTTTCCATTACAGATGACCCCTCAATCCCATCATATAG 1626  
Db 1660 TCTCCTCGAACAACAATTTCTTTCCATTACAGATGACCCCTCAATCCCATCATATAG 1719  
Qy 1627 GTGGGTCCAGAAACCGGCACTACCCCGTTTATTTGGTCTCTCAACATGAGAGAACTC 1686  
Db 1720 GTGGGTCCAGAAACCGGCACTACCCCGTTTATTTGGTCTCTCAACATGAGAGAACTC 1779  
Qy 1687 CAGAAACAACCCAGATGAAATTTTGAAGCAATGTGTTGTTTGGCTGACAGCAT 1746  
Db 1780 CAGAAACAACCCAGATGAAATTTTGAAGCAATGTGTTGTTTGGCTGACAGCAT 1839  
Qy 1747 AAGATGAGGATTTATCTATTCAAGAAAGCTCAGACATTTCTTTAAGCATGGATCTTA 1806  
Db 1840 AAGATGAGGATTTATCTATTCAAGAAAGCTCAGATATTTCTTTAAGCATGGATCTTA 1899  
Qy 1807 ACTCATCTAAAGTTTCTTTCTCAAGAGATGCTCTGTTGGGAGAGGAGGCCAGCA 1866  
Db 1900 ACTCATCTAAAGTTTCTTTCTCAAGAGATGCTCTGTTGGGAGAGGAGGCCAGCA 1959  
Qy 1867 AAGTATGTACAAGACAACATCCAGCTTCAATGGCCAGAGTGGCGAATTCCTCTCAG 1926  
Db 1960 AAGTATGTACAAGACAACATCCAGCTTCAATGGCCAGAGTGGCGAATTCCTCTCAG 2019  
Qy 1927 GAGAACGCGCATATTTATGTGTGTGAGATGCAAGAAATATGGCCAGAGTGTACATGAT 1986  
Db 2020 GAGAACGCGCATATTTATGTGTGTGAGATGCAAGAAATATGGCCAGAGTGTACATGAT 2079  
Qy 1987 GCCCTGTGCAAAATTAAGCAAGAGGTTGGAGTTGAAAACTAAGCAATGAATAAC 2046  
Db 2080 GCCCTGTGCAAAATTAAGCAAGAGGTTGGAGTTGAAAACTAAGCAATGAATAAC 2139  
Qy 2047 CTGGCCACTTTAAAGAGAAACCGCTACCTTCAGAGATATTTGTCTATTA 2097  
Db 2140 CTGGCCACTTTAAAGAGAAACCGCTACCTTCAGAGATATTTGTCTATTA 2190

RESULT 14  
ADQ39030  
ID ADQ39030 standard; DNA; 3274 BP.  
XX  
AC ADQ39030;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Human SNP containing myocardial infarction-associated gene, SEQ ID 693.  
XX  
KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;  
KW cardiac; gene therapy; human; gene; de.  
XX  
OS Homo sapiens.  
XX  
PN NC0204058052-A2.  
XX  
PD 15-JUL-2004.  
XX  
PF 22-DEC-2003; 2003MO-US040978.  
XX  
PR 20-DEC-2002; 2002US-0434778P.  
PR 10-MAR-2003; 2003US-0453135P.  
PR 30-APR-2003; 2003US-046412P.  
PR 23-SEP-2003; 2003US-0504955P.  
XX  
PA (APPL-) APPLERA CORP.  
XX  
PI Cargill M, Devlin J, Iakoubova O;  
XX  
XX WPI; 2004-533949/51.  
DR P-PSDB; ADQ39858.  
XX  
PT Identifying an individual who has an altered risk for developing  
XX myocardial infarction by detecting a single nucleotide polymorphism in



QY 1687 CAAGAACCAACCCAGATGGAATTTTGGAGCAATGTGTTTGGCTGACAGCAT 1746  
CC  
CC 1798 CAGAACCAACCCAGATGGAATTTTGGAGCAATGTGTTTGGCTGACAGCAT 1857  
CC  
QY 1747 AAGGATGGAATTTTCTATTCAGAAAAAGACTCAGACTTTCTTAAAGCATGGGATCTTA 1806  
CC  
CC 1858 AAGGATGGAATTTTCTATTCAGAAAAAGACTCAGACTTTCTTAAAGCATGGGATCTTA 1917  
CC  
QY 1807 ACTCATCTAAAGGTTTCTTCTCAGAGATGCTCTGTTGGGAGAGAGAAAGCCAGCA 1866  
CC  
CC 1918 ACTCATCTAAAGGTTTCTTCTCAGAGATGCTCTGTTGGGAGAGAGAAAGCCAGCA 1977  
CC  
QY 1867 AAGTATGTACAGAACCAACATCCAGCTTCATGCGCAGCAGGTCGAGAACTCTCTCCAG 1926  
CC  
CC 1978 AAGTATGTACAGAACCAACATCCAGCTTCATGCGCAGCAGGTCGAGAACTCTCTCCAG 2037  
CC  
QY 1927 GAGAACGGCCATTTTATGTGTGTGAGATGCAAAAGAAATATGCGCAAGATGACATGAT 1986  
CC  
CC 2038 GAGAACGGCCATTTTATGTGTGTGAGATGCAAAAGAAATATGCGCAAGATGACATGAT 2097  
CC  
QY 1987 GCCCTGTGCAATTAATTAAGCAAAAGAGTTGAGTTGAAAACTAGAACATGAAAAAC 2046  
CC  
CC 2098 GCCCTGTGCAATTAATTAAGCAAAAGAGTTGAGTTGAAAACTAGAACATGAAAAAC 2157  
CC  
QY 2047 CTGGCCACTTTTAAAGAGAAAAACGCTACCTTCAGATATTTGTCATTA 2097  
CC  
CC 2158 CTGGCCACTTTTAAAGAGAAAAACGCTACCTTCAGATATTTGTCATTA 2208  
CC  
Db  
RESULT 15  
ACN42470  
ID ACN42470 standard; cDNA; 3189 BP.  
XX  
XX ACN42470;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
XX Human diagnostic and therapeutic polynucleotide SEQ ID NO:1345.  
XX  
XX ss; gene; gene therapy; human diagnostic and therapeutic polynucleotide;  
XX  
XX dltbp.  
XX  
XX Homo sapiens.  
XX  
XX OS  
XX  
XX MO2004023973-A2.  
XX  
XX PD 25-MAR-2004.  
XX  
XX 12-SEP-2003; 2003WC-US028227.  
XX  
XX 12-SEP-2002; 2002US-0410259P.  
XX  
XX 12-SEP-2002; 2002US-0410260P.  
XX  
XX (INCY-) INCYTE CORP.  
XX  
XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F,  
PI Hartshorne TA, Suchorolski MT, Altus CM, Pites SJ, Elder LV,  
PI Mooney EM, Deleage AM, Panesar IS, Banville SC, Reddy TP,  
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstein EH,  
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve IL,  
PI Lagarde RE, Spito PA, Stewart EA, Wingrove J, Vilt DA, Kirtson ES,  
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D,  
PI Patury S, Shi X, Suarez CJ;  
XX  
XX WPI: 2004-329368/30.  
XX  
XX P-PSDB; ABM83818.  
XX  
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful  
PT in diagnosing a condition, disease or disorder associated with human  
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or  
PT in gene mapping.  
XX  
PS Claim 1, Page: 190pp; English.

XX  
CC The invention relates to novel diagnostic and therapeutic polynucleotides  
CC selected from one of the 2722 sequences defined in the specification. A  
CC polynucleotide of the invention may have a use in gene therapy. The human  
CC diagnostic and therapeutic polynucleotides (dltbp) or polypeptides may be  
CC used to diagnose a particular condition, disease or disorder associated  
CC with human molecules, e.g. cell proliferative disorder, endocrine  
CC autoimmune/inflammatory disorder, developmental disorder, endocrine  
CC disorder, neurological disorders, gastrointestinal disorders, or  
CC infections caused by virus, bacteria, fungi or parasite. The dltbp  
CC molecules may also be used in genetic mapping, in identifying individuals  
CC from minute biological samples, in detecting single nucleotide  
CC polymorphisms, as molecular weight markers, and for somatic or germline  
CC gene therapy. The present sequence represents a dltbp polynucleotide of  
CC the invention. Note: The sequence data for this patent is not represented  
CC in the printed specification, but was obtained in electronic format  
CC directly from WIPO at [www.wipo.int/pct/en/sequences/listing.htm](http://www.wipo.int/pct/en/sequences/listing.htm)  
XX  
SQ Sequence 3189 BP; 916 A; 679 C; 665 G; 929 T; 0 U; 0 Other;  
Query Match 45.6%; Score 956; DB 13; Length 3189;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1056; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ATGAGAGGTTTCTGTTACTATATGCTATACAGACGAGGACGCAAAAGCCATCGCAGAA 60  
Db 112 ATGAGAGGTTTCTGTTACTATATGCTATACAGACGAGGACGCAAAAGCCATCGCAGAA 171  
QY 61 GAAATGTGAGCAAGCTGTGTACATGATGATTTTCTGAGATGCTTACATGATTAAGAA 120  
Db 172 GAAATGTGAGCAAGCTGTGTACATGATGATTTTCTGAGATGCTTACATGATTAAGAA 231  
QY 121 TCCGATAGTATGACCTTAAACCCGAAACAGCTCTCTGTTGTTGTTTACACAG 180  
Db 232 TCCGATAGTATGACCTTAAACCCGAAACAGCTCTCTGTTGTTGTTTACACAG 291  
QY 181 GGCACCGAGAGCCACCCGACACAGCCCGCAAGTTGTTAAGAAATACGAACCAACA 240  
Db 292 GGCACCGAGAGCCACCCGACACAGCCCGCAAGTTGTTAAGAAATACGAACCAACA 351  
QY 241 CTGCGGTTGATTTCTTGTGCTACCTGCGGTATGAGTTACTGAGGTCCTCGGTGATTGAA 300  
Db 352 CTGCGGTTGATTTCTTGTGCTACCTGCGGTATGAGTTACTGAGGTCCTCGGTGATTGAA 411  
QY 301 TACACCTACTTTTGCATGAGGAGGAGATTAATTGATTAACGACTTCAAGACTTGAGCC 360  
Db 412 TACACCTACTTTTGCATGAGGAGGAGATTAATTGATTAACGACTTCAAGACTTGAGCC 471  
QY 361 CGGATTTCTATGACACTGACATGACATGATGATCTGTAGTTTAAACTTGTGTTGAG 420  
Db 472 CGGATTTCTATGACACTGACATGACATGATGATCTGTAGTTTAAACTTGTGTTGAG 531  
QY 421 CGGTGATGCTGAGACTCTGCGCAGCCCTCAGAAAGCAATTTAGTCAAGAGAGCAA 480  
Db 532 CGGTGATGCTGAGACTCTGCGCAGCCCTCAGAAAGCAATTTAGTCAAGAGAGCAA 591  
QY 481 GAGAGATTAAGTGGCGCACTCCGATGACATCACTGACATCTTGAAGACAGACTTGTG 540  
Db 592 GAGAGATTAAGTGGCGCACTCCGATGACATCACTGACATCTTGAAGACAGACTTGTG 651  
QY 541 AAGTCAGAGCTGCTACATGATTAATCTCAAGTGAAGTTTGAATTCATGATTCAGGA 600  
Db 652 AAGTCAGAGCTGCTACATGATTAATCTCAAGTGAAGTTTGAATTCATGATTCAGGA 711  
QY 601 AGAAGGATCTGAGGTTTGAAGCAAAATAGCAGTGAACAGCAACCAATCAATGTTGA 660  
Db 712 AGAAGGATCTGAGGTTTGAAGCAAAATAGCAGTGAACAGCAACCAATCAATGTTGA 771  
QY 661 ATTGAAGACTTTAGTCTCTCACTTACCCGTTGAGTACCCCACTGACAGAGCTCTCTG 720  
Db 772 ATTGAAGACTTTAGTCTCTCACTTACCCGTTGAGTACCCCACTGACAGAGCTCTCTG 831  
QY 721 AATATTCCTGTTTACCCCAAGATATTTACAGGTACATCTGACGAGAGTCTCTTGGCAG 780

```
Db      832 AATATTCCTGTTTACCCCGAATATTACAGTACATCTGCAGAGTCTCTGGCCAG 891
Qy      781 GAGGAAGCCAGTATCTGTGACTTCAGAGATCCAGTTTCAAGTGCATTTCAAG 840
Db      892 GAGGAAGCCAGTATCTGTGACTTCAGAGATCCAGTTTCAAGTGCATTTCAAG 951
Qy      841 GCAGTTCAACTTACTAGATGATGCCATAAAAACCACTCTGCTGTAGATTGACATT 900
Db      952 GCAGTTCAACTTACTAGATGATGCCATAAAAACCACTCTGCTGTAGATTGACATT 1011
Qy      901 TCAAAATACAGACTTTTCTATCAGCCTGAGATGCCCTTCAGCGTGAATCTGCCCTAACAGT 960
Db      1012 TCAAAATACAGACTTTTCTATCAGCCTGAGATGCCCTTCAGCGTGAATCTGCCCTAACAGT 1071
Qy      961 GATTCTGAGGTACAAAGCCTACTCCAAAGACTGACCTTGAAATTAAGAGACACTGC 1020
Db      1072 GATTCTGAGGTACAAAGCCTACTCCAAAGACTGACCTTGAAATTAAGAGACACTGC 1131
Qy      1021 GTCCCTTTGAAATTAAGGCAGACACAAGAGAAAGG 1058
Db      1132 GTCCCTTTGAAATTAAGGCAGACACAAGAGAAAGG 1169
```

Search completed: August 27, 2005, 01:18:26  
Job time : 742.49 secs



***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 23:30:25 ; Search time 235.757 Seconds  
(without alignments)  
14554.251 Million cell updates/sec

Title: US-09-371-347A-1

Perfect score: 2097

Sequence: 1 atgaggaggttcctgctact.....ttccagatattggtcctaa 2097

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq:\*
- 2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq:\*
- 3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq:\*
- 4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*
- 5: /cgn2\_6/prodata/1/ina/6C\_COMB.seq:\*
- 6: /cgn2\_6/prodata/1/ina/6D\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2097	100.0	3259	3 US-09-318-448-23	Sequence 23, Appl
2	1944	92.7	3242	4 US-09-949-016-4215	Sequence 4215, Ap
3	386	18.4	390	3 US-08-905-223-71	Sequence 71, Appl
4	330	15.7	601	4 US-09-849-016-150019	Sequence 150019, A
5	279	13.3	35916	4 US-09-849-016-150020	Sequence 150020, A
6	279	13.3	601	4 US-09-949-016-150037	Sequence 150037, A
7	189	9.0	601	4 US-09-949-016-150030	Sequence 150030, A
8	189	9.0	2475	4 US-09-949-016-150030	Sequence 150030, A
9	155	7.4	601	4 US-09-849-016-150031	Sequence 150031, A
10	145	6.9	601	4 US-09-849-016-150046	Sequence 150046, A
11	137	6.5	601	4 US-09-849-016-150047	Sequence 150047, A
12	137	6.5	601	4 US-09-849-016-150029	Sequence 150029, A
13	125	6.0	601	4 US-09-949-016-150041	Sequence 150041, A
14	121	5.8	601	4 US-09-949-016-150042	Sequence 150042, A
15	121	5.8	601	4 US-09-849-016-150008	Sequence 150008, A
16	119	5.7	601	4 US-09-849-016-150055	Sequence 150055, A
17	119	5.7	601	4 US-09-849-016-150048	Sequence 150048, A
18	110	5.2	601	4 US-09-949-016-150032	Sequence 150032, A
19	94	4.5	601	4 US-09-471-276-495	Sequence 495, App
20	78	3.7	244	4 US-09-849-016-150018	Sequence 150018, A
21	78	3.7	601	4 US-09-849-016-150018	Sequence 150018, A
22	76	3.6	601	4 US-09-849-016-150018	Sequence 150018, A
23	30	1.4	1681	4 US-09-849-016-150018	Sequence 150018, A
24	20	1.0	273	4 US-09-849-016-150018	Sequence 150018, A
25	20	1.0	440	3 US-09-397-787-305	Sequence 305, App
26	20	1.0	444	4 US-09-397-787-305	Sequence 305, App
27	20	1.0	445	3 US-09-397-787-274	Sequence 274, App

C 28	20	1.0	174259	4 US-09-949-016-11968	Sequence 11968, A
C 29	20	1.0	174262	4 US-09-949-016-14259	Sequence 14259, A
C 30	19	0.9	169	1 US-08-166-346A-8	Sequence 8, Appl
C 31	19	0.9	459	4 US-09-621-978-8324	Sequence 8324, Ap
C 32	19	0.9	3969	3 US-09-518-386B-4	Sequence 4, Appl
C 33	19	0.9	4396	3 US-09-821-736-1	Sequence 1, Appl
C 34	19	0.9	14721	4 US-09-949-016-13507	Sequence 13507, A
C 35	19	0.9	25199	4 US-09-949-016-13361	Sequence 13361, A
C 36	19	0.9	129658	4 US-09-949-016-17195	Sequence 17195, A
C 37	19	0.9	186734	4 US-09-949-016-14870	Sequence 14870, A
C 38	19	0.9	193689	4 US-09-949-016-12350	Sequence 12350, A
C 39	19	0.9	193689	4 US-09-949-016-13088	Sequence 13088, A
C 40	19	0.9	200663	4 US-09-949-016-12569	Sequence 12569, A
C 41	19	0.9	203093	4 US-09-949-016-14445	Sequence 14445, A
C 42	18	0.9	78	2 US-08-749-852-56	Sequence 56, Appl
C 43	18	0.9	78	2 US-08-749-852-58	Sequence 58, Appl
C 44	18	0.9	511	4 US-09-902-540-1374	Sequence 1374, Ap
C 45	18	0.9	531	4 US-09-252-991A-2223	Sequence 2223, Ap

## ALIGNMENTS

RESULT 1					
US-09-318-448-23					
; Sequence 23, Application US/09318448					
; Patent No. 6210950					
; GENERAL INFORMATION:					
; APPLICANT: Johnson, William G.					
; TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING					
; FILE REFERENCE: 601-1-057					
; CURRENT APPLICATION NUMBER: US/09/318,448					
; NUMBER OF SEQ ID NOS: 46					
; SOFTWARE: PatentIn Ver. 2.0					
; SEQ ID NO 23					
; LENGTH: 3259					
; TYPE: DNA					
; ORGANISM: Homo sapiens					
US-09-318-448-23					
Query Match					
Best Local Similarity 100.0%; Score 2097; DB 3; Length 3259;					
Matches 2097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	ATGAGGAGGTTCTGTACTATATCTACACAGCAGGAGGAGCAAGGCCATCGCAGAA	60		
DB	80	ATGAGGAGGTTCTGTACTATATCTACACAGCAGGAGGAGCAAGGCCATCGCAGAA	139		
QY	61	GAATGTGTGACCAAGCTGTGTACATGATTTTCTGCAGATCTTCACTGTAATGTA	120		
DB	140	GAATGTGTGACCAAGCTGTGTACATGATTTTCTGCAGATCTTCACTGTAATGTA	199		
QY	121	TCCGTAAGTATGACCTTAAACCGAAACAGCTCTCTGTGTGTGTCTTCTACAG	180		
DB	200	TCCGTAAGTATGACCTTAAACCGAAACAGCTCTCTGTGTGTGTCTTCTACAG	259		
QY	181	GGCAGCGAGAGCCACCCAGACAGCCGCAAGTTGTTAAGAAATACAGAACCA	240		
DB	260	GGCAGCGAGAGCCACCCAGACAGCCGCAAGTTGTTAAGAAATACAGAACCA	319		
QY	241	CTGCGGTTGATTTCTTGTCTACCTGCGGTAAGGTTAAGGTTCTCGGTATTCAG	300		
DB	320	CTGCGGTTGATTTCTTGTCTACCTGCGGTTAAGGTTAAGGTTCTCGGTATTCAG	379		
QY	301	TACACTTATTTTGAATGGGGGGAATTAATTAAGCACTTAAGCTTGAGGC	360		
DB	380	TACACTTATTTTGAATGGGGGGAATTAATTAAGCACTTAAGCTTGAGGC	439		
QY	361	CGGATTTTATGACATGACATGATGATGATGATGATGATGATGATGATGATGAT	420		

Dh 440 CGGCAATTTCTATGACACTGACATGACATGACTGTGTAGGTTTGAACAATTGGTTGAG 449  
Qy 421 CCGTGATGCTGGACTCTGGCCAGCCCTCAGAAAGCAATTTAGTCAACAGAGACA 480  
Dh 500 CCGTGATGCTGGACTCTGGCCAGCCCTCAGAAAGCAATTTAGTCAACAGAGACA 559  
Qy 481 GAGAGATTAAGTGGCGCACTCCGGTGGCATCACCTGATCTTTGAGAGACAGCTTTGTG 540  
Dh 560 GAGAGATTAAGTGGCGCACTCCGGTGGCATCACCTGATCTTTGAGAGACAGCTTTGTG 619  
Qy 541 AAGTCAGAGCTGCTACATTTGAATCTCAAGTGGAGCTTGGAGATTCAGATTCAGGA 600  
Dh 620 AAGTCAGAGCTGCTACATTTGAATCTCAAGTGGAGCTTGGAGATTCAGATTCAGGA 679  
Qy 601 AGAAGGATTCAGAGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATGTTGTA 660  
Dh 680 AGAAGGATTCAGAGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATGTTGTA 739  
Qy 661 ATTGAAGATTTGAGTCTCTCACTTACCCTTGGTACCCCACTCTCAACAGCTCTCTG 720  
Dh 740 ATTGAAGATTTGAGTCTCTCACTTACCCTTGGTACCCCACTCTCAACAGCTCTCTG 799  
Qy 721 AATATTCCTGGTTAACCCCGAGATATTTTACAGGTACATGACAGAGTCTCTGGCCAG 780  
Dh 800 AATATTCCTGGTTAACCCCGAGATATTTTACAGGTACATGACAGAGTCTCTGGCCAG 859  
Qy 781 GAGGAAAGCCAAAGTATCTGTGACTTTCAGCAGATCAGATTTTTCAGTGCCAAATTTCAA 840  
Dh 860 GAGGAAAGCCAAAGTATCTGTGACTTTCAGCAGATCAGATTTTTCAGTGCCAAATTTCAA 919  
Qy 841 GCAGTTCAATTTAGTATGATGATGCAATTAACCACTGCTGTGTGAATTTGACATTT 900  
Dh 920 GCAGTTCAATTTAGTATGATGATGCAATTAACCACTGCTGTGTGAATTTGACATTT 979  
Qy 901 TCAATTCAGACTTTTCTTCAATCAGCTGAGATGCTTCAAGCTGATCTGCTTCAACAT 960  
Dh 980 TCAATTCAGACTTTTCTTCAATCAGCTGAGATGCTTCAAGCTGATCTGCTTCAACAT 1039  
Qy 961 GATTTCTAGAGTACAAAGCTATCTCAAAAGCTGAGCTTGAAGTAAAGAGACACTGC 1020  
Dh 1040 GATTTCTAGAGTACAAAGCTATCTCAAAAGCTGAGCTTGAAGTAAAGAGACACTGC 1099  
Qy 1021 GTCTTTTGAATTAAGGCGAGACACAAAGAAAGAGCTACTTACCAGCATATA 1080  
Dh 1100 GTCTTTTGAATTAAGGCGAGACACAAAGAAAGAGCTACTTACCAGCATATA 1159  
Qy 1081 CCTGCGGAGATGTTCTCTCAGATTCATTTTACCTGTGTCTGAAATCCGAGCAATTCCT 1140  
Dh 1160 CCTGCGGAGATGTTCTCTCAGATTCATTTTACCTGTGTCTGAAATCCGAGCAATTCCT 1219  
Qy 1141 AAAAAAGCATTTTTCGAGCCCTTGTGGAATTAACAAGTGAAGTGTGAAAAGCGCAGG 1200  
Dh 1220 AAAAAAGCATTTTTCGAGCCCTTGTGGAATTAACAAGTGAAGTGTGAAAAGCGCAGG 1279  
Qy 1201 CTAAGAGAGCTGTGCAATTAACAAGGCGCAGCCGATTAATAGCCGCTTTTGTACAGAGTCC 1260  
Dh 1280 CTAAGAGAGCTGTGCAATTAACAAGGCGCAGCCGATTAATAGCCGCTTTTGTACAGAGTCC 1339  
Qy 1261 TGTGCTGCTGTTGGAATCTCTCTGCTTCCCTTTCGAGCCAGCACTGAGTCTC 1320  
Dh 1340 TGTGCTGCTGTTGGAATCTCTCTGCTTCCCTTTCGAGCCAGCACTGAGTCTC 1399  
Qy 1321 CTGCTCGAACAATCTTCTTAACCTTCAACCCAGACCATATTCGTGTGCAAGCTCAAGTTTA 1380  
Dh 1400 CTGCTCGAACAATCTTCTTAACCTTCAACCCAGACCATATTCGTGTGCAAGCTCAAGTTTA 1459  
Qy 1381 TTTTCAACCAAGAAAGCTCTCAATTTTGTCTTCAACATTTGTGCAATTTTGTCTACTGCCACA 1440  
Dh 1460 TTTTCAACCAAGAAAGCTCTCAATTTTGTCTTCAACATTTGTGCAATTTTGTCTACTGCCACA 1519  
Qy 1441 ACAGAGGTTCTGCGAAGGAGATGTAAGAGGCTGAGCTGCTTGTGTTGTTGCTTCAAGTT 1500  
Dh 1520 ACAGAGGTTCTGCGAAGGAGATGTAAGAGGCTGAGCTGCTTGTGTTGTTGCTTCAAGTT 1579

Qy 1501 CTTGAGCCAAACATACATGATCCATGAAAGACGGGGAAGACCTTGCTCTTAAGATA 1560  
Dh 1580 CTTGAGCCAAACATACATGATCCATGAAAGACGGGGAAGACCTTGCTCTTAAGATA 1639  
Qy 1561 TCCATCTCTCTGGAACAACAATTTCTTTCACATTAACAGATGACCCCTCAATCCCATC 1620  
Dh 1640 TCCATCTCTCTGGAACAACAATTTCTTTCACATTAACAGATGACCCCTCAATCCCATC 1699  
Qy 1621 ATAAATGAGGATCCAGAAACCGGCAATAGCCGCTTATTTAGGTTCTTACAACTAGAGAG 1680  
Dh 1700 ATAAATGAGGATCCAGAAACCGGCAATAGCCGCTTATTTAGGTTCTTACAACTAGAGAG 1759  
Qy 1681 AAATCTCAAGAAACAACCCAGATGAAATTTTGAAGCAATGTGTTTGTGCTGC 1740  
Dh 1760 AAATCTCAAGAAACAACCCAGATGAAATTTTGAAGCAATGTGTTTGTGCTGC 1819  
Qy 1741 AGGCATTAAGATGGAATTAATCTAATTCAGAAAGACCTCAGATTTCTTAAGCATGGG 1800  
Dh 1820 AGGCATTAAGATGGAATTAATCTAATTCAGAAAGACCTCAGATTTCTTAAGCATGGG 1879  
Qy 1801 ATCTTAATCTAATTAAGGTTTCTTCTCAAGAGATGCTCTGTGGGAGAGAGAGCC 1860  
Dh 1880 ATCTTAATCTAATTAAGGTTTCTTCTCAAGAGATGCTCTGTGGGAGAGAGAGCC 1939  
Qy 1861 CCAAGCAAGTATGTAACAACAATCCAGCTTCAATGCGCAGAGTGGCGAGATCTTC 1920  
Dh 1940 CCAAGCAAGTATGTAACAACAATCCAGCTTCAATGCGCAGAGTGGCGAGATCTTC 1999  
Qy 1921 CTCAGAGAAACGGCCATATTTATGTGTGAGATGAGCAAGAAATATGCGCAAGATGTA 1980  
Dh 2000 CTCAGAGAAACGGCCATATTTATGTGTGAGATGAGCAAGAAATATGCGCAAGATGTA 2059  
Qy 1981 CATGATGCCCTTGTGCAAAATTAATAGCAAGAAGTTGAGTTGAAAAACTAGAAAGCATG 2040  
Dh 2060 CATGATGCCCTTGTGCAAAATTAATAGCAAGAAGTTGAGTTGAAAAACTAGAAAGCATG 2119  
Qy 2041 AAAACCTGCGCACTTTAAAGAAAGAAAGCGTACCTTCAGAGATTTTGTCTATA 2097  
Dh 2120 AAAACCTGCGCACTTTAAAGAAAGAAAGCGTACCTTCAGAGATTTTGTCTATA 2176

RESULT 2  
US-09-949-016-4215  
; Sequence 4215, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTNER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4215  
; LENGTH: 3242  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-4215

Query Match 92.7%; Score 1944; DB 4; Length 3242;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2094; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGAGAGGTTTCTGTTACTATATGCTACACAGAGGAAAGGCGCATGCGAGAA 60  
|||||

80 ATGAGAGGTTCTGTACTATATGCTACACAGACAGGACAGGCAAGGCCATCCGACAA 139  
61 GAAATGTGACCAAGCTGTGTACATGATTTCTGAGATCTTCACTGATTAAGTGA 120  
140 GAAATATGTGACCAAGCTGTGTACATGATTTCTGAGATCTTCACTGATTAAGTGA 199  
121 TCCGATTAAGTATGACCTAAACCCGAAACAGCTCTCTGTTGTTGTGTTCTACACG 180  
200 TCCGATTAAGTATGACCTAAACCCGAAACAGCTCTCTGTTGTTGTGTTCTACACG 259  
181 GGCACCGGAGACCCGACACAGCCGCAAGTTGTTAAGAAATACAAACCAACA 240  
260 GGCACCGGAGACCCGACACAGCCGCAAGTTGTTAAGAAATACAAACCAACA 319  
241 CTGCGGTGATTTCTGTGCTCACTGCGGTATGAGTTAGTGGTCTGCTGATTCAGAA 300  
320 CTGCGGTGATTTCTGTGCTCACTGCGGTATGAGTTAGTGGTCTGCTGATTCAGAA 379  
301 TACACCTACTTTTGGCAATGGGGGGAAGATTAATGATTAACGACTTCAGAGCTTGAGCC 360  
380 TACACCTACTTTTGGCAATGGGGGGAAGATTAATGATTAACGACTTCAGAGCTTGAGCC 439  
361 CGGCAATTTCTATGACATGACATGACATGACATGACATGACATGACATGACATGAC 420  
440 CGGCAATTTCTATGACATGACATGACATGACATGACATGACATGACATGACATGAC 499  
421 CGGTGATTTGCTGACATGACATGACATGACATGACATGACATGACATGACATGAC 480  
500 CGGTGATTTGCTGACATGACATGACATGACATGACATGACATGACATGACATGAC 559  
481 GAGAGATTAAGTGTGACATGACATGACATGACATGACATGACATGACATGACATGAC 540  
560 GAGAGATTAAGTGTGACATGACATGACATGACATGACATGACATGACATGACATGAC 619  
541 AAGTCAGAGCTGTGACATGACATGACATGACATGACATGACATGACATGACATGAC 600  
620 AAGTCAGAGCTGTGACATGACATGACATGACATGACATGACATGACATGACATGAC 679  
601 AAGAGATTTCTGAGTTTGAAGCAAAATGACATGACATGACATGACATGACATGACAT 660  
680 AAGAGATTTCTGAGTTTGAAGCAAAATGACATGACATGACATGACATGACATGACAT 739  
661 ATTGAAGCTTTGAGCTCACTTACCCGTTGCGGTACCCGCTCTCAAGGCTCTCTG 720  
740 ATTGAAGCTTTGAGCTCACTTACCCGTTGCGGTACCCGCTCTCTCAAGGCTCTCTG 799  
721 AATATTTCTGTGTTTACCCGCAAAATTAATGACATGACATGACATGACATGACATGAC 780  
800 AATATTTCTGTGTTTACCCGCAAAATTAATGACATGACATGACATGACATGACATGAC 859  
781 GAGGAAAGCCAGATCTGTGACTTACGACATGACATGACATGACATGACATGACATGAC 840  
860 GAGGAAAGCCAGATCTGTGACTTACGACATGACATGACATGACATGACATGACATGAC 919  
841 GAGGAAAGCCAGATCTGTGACTTACGACATGACATGACATGACATGACATGACATGAC 900  
920 GAGGAAAGCCAGATCTGTGACTTACGACATGACATGACATGACATGACATGACATGAC 979  
901 TCAATATACAGACTTTCTATCAGCTGAGATGCTTCAAGCTGTGCTTCAAGT 960  
980 TCAATATACAGACTTTCTATCAGCTGAGATGCTTCAAGCTGTGCTTCAAGT 1039  
961 GATTTGAGGTACAAAGCTTCTCAAGATGCTGAGCTTGAAGTAAAGAGACCTGC 1020  
1040 GATTTGAGGTACAAAGCTTCTCAAGATGCTGAGCTTGAAGTAAAGAGACCTGC 1099  
1021 GTTCCTTTGAAATTAAGGACACAAAGAAAGAGGCTTACCTTACCCGACATTA 1080  
1100 GTTCCTTTGAAATTAAGGACACAAAGAAAGAGGCTTACCTTACCCGACATTA 1159  
1081 CTGTGCGGATGTTCTCTCAGATTCATTTTCTGCTGCTTGAAGTCCGAGCAATTCCT 1140  
1160 CTGTGCGGATGTTCTCTCAGATTCATTTTCTGCTGCTTGAAGTCCGAGCAATTCCT 1219

1141 AAAAAGCATTTTTCGAGCCCTTGTGACATATACAGTACAGTCTGAAAAAGCGACG 1200  
1220 AAAAAGCATTTTTCGAGCCCTTGTGACATATACAGTACAGTCTGAAAAAGCGACG 1279  
1201 CTACAGAGCTGTGACATTAACAGAGGAGCGCATTAATAGCCCTTGTACAGATGCTC 1260  
1280 CTACAGAGCTGTGACATTAACAGAGGAGCGCATTAATAGCCCTTGTACAGATGCTC 1339  
1261 TGTGCTGCTGTGTGATCT 1320  
1340 TGTGCTGCTGTGTGATCT 1399  
1321 CTGCTGAAACATCTTCTTAACTTCAACCCAGACCATATTCGTGTGCAAGTCAAGTTTA 1380  
1400 CTGCTGAAACATCTTCTTAACTTCAACCCAGACCATATTCGTGTGCAAGTCAAGTTTA 1459  
1381 TTTCAACCCAGAAAGCTCATTTTGTCTTCAACATTTGTGAATTTCTGTCTACGACAA 1440  
1460 TTTCAACCCAGAAAGCTCATTTTGTCTTCAACATTTGTGAATTTCTGTCTACGACAA 1519  
1441 ACAGAGTTCTGCGGAAGGAGATATGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
1520 ACAGAGTTCTGCGGAAGGAGATATGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1579  
1501 CTTCAGCCCAACATACATGACATCCCATGAAACAGCGGAAAGCCCTGCTCTTAAGATA 1560  
1580 CTTCAGCCCAACATACATGACATCCCATGAAACAGCGGAAAGCCCTGCTCTTAAGATA 1639  
1561 TCCATCTCTCTCTGACACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
1640 TCCATCTCTCTCTGACACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1699  
1621 ATTAATGTGTGTGTCAGAAACCGGCAATGCGCTTATTTGGTCTCTTACAAATAGAG 1680  
1700 ATTAATGTGTGTGTCAGAAACCGGCAATGCGCTTATTTGGTCTCTTACAAATAGAG 1759  
1681 AAATCTCAAGAACCAACCCAGATGAAATTTTGGAGCAATGTGTTTGTGCTGCTC 1740  
1760 AAATCTCAAGAACCAACCCAGATGAAATTTTGGAGCAATGTGTTTGTGCTGCTC 1819  
1741 AGGCAATTAAGATTAAGATTAATCTATTCAGAAAGAGCTCAGACATTTCTTAAAGATG 1800  
1820 AGGCAATTAAGATTAAGATTAATCTATTCAGAAAGAGCTCAGACATTTCTTAAAGATG 1879  
1801 ATCTTAATCTATTAAGATTTCTTCTTAAAGATGCTCTGTTGGGAGAGAGAGCC 1860  
1880 ATCTTAATCTATTAAGATTTCTTCTTAAAGATGCTCTGTTGGGAGAGAGAGCC 1939  
1861 CCAGCAAGATTAATTAAGCAACATCCAGCTTCAATGCGCAGAGGTGCGAGAAATCTC 1920  
1940 CCAGCAAGATTAATTAAGCAACATCCAGCTTCAATGCGCAGAGGTGCGAGAAATCTC 1999  
1921 CTTCAGAGAGAGGCGCATTTATGTGTGTGAGATGCAAGAAATATGCGCAAGATGTA 1980  
2000 CTTCAGAGAGAGGCGCATTTATGTGTGTGAGATGCAAGAAATATGCGCAAGATGTA 2059  
1981 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGGTTGAGTTGAAAACTTAAAGCAATG 2040  
2060 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGGTTGAGTTGAAAACTTAAAGCAATG 2119  
2041 AAAAACCCTGCGCACTTTTAAAGAAAGAAAGCCCTCAGAGATTTTGTGATTA 2097  
2120 AAAAACCCTGCGCACTTTTAAAGAAAGAAAGCCCTCAGAGATTTTGTGATTA 2176

RESULT 3  
US-08-905-223-71  
; Sequence 71, Application US/08905223  
; Patent No. 6222029  
; GENERAL INFORMATION:  
; APPLICANT: Edwards, Jean-Baptiste D.  
; APPLICANT: Duclercq, Aymeric

APPLICANT: Lacroix, Bruno  
TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS  
NUMBER OF SEQUENCES: 503  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe, Martens, Olson & Bear  
STREET: 501 West Broadway  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92101-3505  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: win95  
SOFTWARE: word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/905,223  
FILING DATE:  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Israel, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 71:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 390 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: DOUBLE  
TOPOLOGY: LINEAR  
MOLECULE TYPE: CDNA  
ORIGINAL SOURCE:  
ORGANISM: Homo Sapiens  
TISSUE TYPE: Brain  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 289..357  
IDENTIFICATION METHOD: Von Heijne matrix  
OTHER INFORMATION: score 6.9  
OTHER INFORMATION: seq SLSLSLASHSVSC/SN  
US-08-905-223-71

Query Match 18.4%; Score 386; DB 3; Length 390;  
Best Local Similarity 100.0%; Pred. No. 3.5e-188;  
Matches 386; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 970 GTACAAAGCTCTCTCAAGAGCTGAGCTTGAAGATAAAGAGAGCACTGCTCTTTTG 1029  
DB 3 GTACAAAGCTCTCTCAAGAGCTGAGCTTGAAGATAAAGAGAGCACTGCTCTTTTG 62  
QY 1030 AAAATAAGGAGACACAAAGAGAGAGCTACCTTACCCGACATATACCTGGGGA 1089  
DB 63 AAAATAAGGAGACACAAAGAGAGAGCTACCTTACCCGACATATACCTGGGGA 122  
QY 1090 TGTCTCTCCAGTTCATTTTACCTGCTGTCTTGAATCCGAGCAATTCCTAAAAAGCA 1149  
DB 123 TGTCTCTCCAGTTCATTTTACCTGCTGTCTTGAATCCGAGCAATTCCTAAAAAGCA 182  
QY 1150 TTTTGGAGAGCCCTTGGAGATATACAGTACAGTGTGTAAAAAGCCAGGCTACAGAG 1209  
DB 183 TTTTGGAGAGCCCTTGGAGATATACAGTACAGTGTGTAAAAAGCCAGGCTACAGAG 242  
QY 1210 CTGTCAGTAAACAAGGGGAGCGGATATAGCGGCTTTGTAAGAGATGCTGAGCTGC 1269  
DB 243 CTGTCAGTAAACAAGGGGAGCGGATATAGCGGCTTTGTAAGAGATGCTGAGCTGC 302  
QY 1270 TTGTGGATCTGAA 1329  
DB 303 TTGTGGATCTGAA 362  
QY 1330 CATCTCTTAACCTTCAACCCAGAC 1355

DB 363 CATCTCTTAACCTTCAACCCAGAC 388

RESULT 4  
US-09-949-016-150019  
Sequence 150019, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 150019  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-150019

Query Match 15.7%; Score 330; DB 4; Length 601;  
Best Local Similarity 99.7%; Pred. No. 2.5e-159;  
Matches 380; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 401 GTTTAGACTTGTGTGAGCCCTGTGATGCTGAGCTTGGCCAGCCCTCAGAAACATT 460  
DB 178 GTTTAGACTTGTGTGAGCCCTGTGATGCTGAGCTTGGCCAGCCCTCAGAAACATT 237  
QY 461 TTAGTCAAGCAGAGACAAAGAGATTAATGCGCATCTCCGGTGCATCCTGCAT 520  
DB 238 TTAGTCAAGCAGAGACAAAGAGATTAATGCGCATCTCCGGTGCATCCTGCAT 297  
QY 521 CTTGAGGACAGACCTTGAAGTCAAGTGTACATGTAATCAAGTGAAGCTTC 580  
DB 298 CTTGAGGACAGACCTTGAAGTCAAGTGTACATGTAATCAAGTGAAGCTTC 357  
QY 581 TGAGATTCGATGATTCAGAGAAAGAGATTCGAGGTTTGAAGCAAAATGACGTAC 640  
DB 358 TGAGATTCGATGATTCAGAGAAAGAGATTCGAGGTTTGAAGCAAAATGACGTAC 417  
QY 641 GCAACCAATCAATGTTGTAATGAAGACTTGAAGTCTTACCTTACCCGATATTTACAGATCATC 700  
DB 418 GCAACCAATCAATGTTGTAATGAAGACTTGAAGTCTTACCTTACCCGATATTTACAGATCATC 477  
QY 701 CACTCTCAGAGCTCTCTGAATATTCCTGTTTACCCCGAGATATTTACAGATCATC 760  
DB 478 CACTCTCAGAGCTCTCTGAATATTCCTGTTTACCCCGAGATATTTACAGATCATC 537  
QY 761 TGCAGAGTCTCTTGGCCAGG 781  
DB 538 TGCAGAGTCTCTTGGCCAGG 558

RESULT 5  
US-09-949-016-159957  
Sequence 159957, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 15957  
LENGTH: 35916  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-15957

Query Match 15.7%; Score 330; DB 4; Length 35916;  
Best Local Similarity 99.7%; Pred. No. 3.1e-159;  
Matches 380; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 401 GTTTAGAACTTGTGTGAGCCGTGATTCCTGACCTGCGCAGCCCTCAGAAAGCATT 460  
DB 10781 GTTTAGAACTTGTGTGAGCCGTGATTCCTGACCTGCGCAGCCCTCAGAAAGCATT 10840  
QY 461 TTAGGTCAAGCAGAGCAAGAGAGATTAAGTGGCGCACTCCGGTGGCATCCTGCAT 520  
DB 10841 TTAGGTCAAGCAGAGCAAGAGAGATTAAGTGGCGCACTCCGGTGGCATCCTGCAT 10900  
QY 521 CCTTGAAGCAGACCTTGTGAAGTCAAGCTGCTACACATTAATCTCAAGTGCAGCTTC 580  
DB 10901 CCTTGAAGCAGACCTTGTGAAGTCAAGCTGCTACACATTAATCTCAAGTGCAGCTTC 10960  
QY 581 TGAGATTGCATGATTCAGAGAAAGAGATTCAGAGTTTGAAGCAAAATGACGTAAACA 640  
DB 10961 TGAGATTGCATGATTCAGAGAAAGAGATTCAGAGTTTGAAGCAAAATGACGTAAACA 11020  
QY 641 GCAACCAATCCAAATGTTGATTAATGAAGCTTGAAGTCTCACTTACCCTTGGTACCCC 700  
DB 11021 GCAACCAATCCAAATGTTGATTAATGAAGCTTGAAGTCTCACTTACCCTTGGTACCCC 11080  
QY 701 CACTTCACAAGCTCTCTGAATATTCCTGTTTACCCCAAGAAATTTTACAGGTACATC 760  
DB 11081 CACTTCACAAGCTCTCTGAATATTCCTGTTTACCCCAAGAAATTTTACAGGTACATC 11140  
QY 761 TCGAGAGTCTCTTGGCCAGG 781  
DB 11141 TCGAGAGTCTCTTGGCCAGG 11161

## RESULT 6

US-09-949-016-150020  
Sequence 150020, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 150020  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-150020

Query Match 13.3%; Score 279; DB 4; Length 601;

Best Local Similarity 99.5%; Pred. No. 4.6e-133;  
Matches 379; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 401 GTTTAGAACTTGTGTGAGCCGTGATTCCTGACCTGCGCAGCCCTCAGAAAGCATT 460  
DB 165 GTTTAGAACTTGTGTGAGCCGTGATTCCTGACCTGCGCAGCCCTCAGAAAGCATT 224  
QY 461 TTAGGTCAAGCAGAGCAAGAGAGATTAAGTGGCGCACTCCGGTGGCATCCTGCAT 520  
DB 225 TTAGGTCAAGCAGAGCAAGAGAGATTAAGTGGCGCACTCCGGTGGCATCCTGCAT 284  
QY 521 CCTTGAAGCAGACCTTGTGAAGTCAAGCTGCTACACATTAATCTCAAGTGCAGCTTC 580  
DB 285 CCTTGAAGCAGACCTTGTGAAGTCAAGCTGCTACACATTAATCTCAAGTGCAGCTTC 344  
QY 581 TGAGATTGCATGATTCAGAGAAAGAGATTCAGAGTTTGAAGCAAAATGACGTAAACA 640  
DB 345 TGAGATTGCATGATTCAGAGAAAGAGATTCAGAGTTTGAAGCAAAATGACGTAAACA 404  
QY 641 GCAACCAATCCAAATGTTGATTAATGAAGCTTGAAGTCTCACTTACCCTTGGTACCCC 700  
DB 405 GCAACCAATCCAAATGTTGATTAATGAAGCTTGAAGTCTCACTTACCCTTGGTACCCC 464  
QY 701 CACTTCACAAGCTCTCTGAATATTCCTGTTTACCCCAAGAAATTTTACAGGTACATC 760  
DB 465 CACTTCACAAGCTCTCTGAATATTCCTGTTTACCCCAAGAAATTTTACAGGTACATC 524  
QY 761 TCGAGAGTCTCTTGGCCAGG 781  
DB 525 TCGAGAGTCTCTTGGCCAGG 545

## RESULT 7

US-09-949-016-150037  
Sequence 150037, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 150037  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-150037

Query Match 9.0%; Score 189; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 1e-86;  
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1369 AGCTCAAGTTATTTTCAACCAAGAAAGCTCCATTTTGTCTTCAACATTTGGAATTTCTG 1428  
DB 18 AGCTCAAGTTATTTTCAACCAAGAAAGCTCCATTTTGTCTTCAACATTTGGAATTTCTG 77  
QY 1429 TCTACTGCCACAAGAGGTTCTGCGAAGGAGATATGTACAGGCTGCTGCGCTTGTG 1488  
DB 78 TCTACTGCCACAAGAGGTTCTGCGAAGGAGATATGTACAGGCTGCTGCGCTTGTG 137  
QY 1489 GTTGTCTCAGTTCTTCAAGCAAAATATCATGATCCATGAAGAAGCGGGAAGCCCTG 1548  
DB 138 GTTGTCTCAGTTCTTCAAGCAAAATATCATGATCCATGAAGAAGCGGGAAGCCCTG 197

QY	1549	GCTCCTAAG	1557
Db	198	GCTCCTAAG	206

RESULT 8  
US-09-566-921-88  
; Sequence 88, Application US/09566921  
Datafile: C00000

OTHER INFORMATION: a, t, c, g, or other  
US-09-566-921-88

```

RESULT 9
US-09-949-016-150030
; Sequence 150030, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241, 755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237, 768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231, 498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 150030
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-150030

```

Query Match	7.4%;	Score 155;	DB 4;	Length 601;
Best Local Similarity	100.0%;	Pred. No. 3.4e-69;		
Matches 155;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```

RESULT 10
US-09-949-016-150031
; Sequence 150031, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,458
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 150031
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-150031

```

Query Match	6.9%	Score 145	DB 4	Length 601
Best Local Similarity	100.0%	Pred. No. 4.8e-64		
Matches 145; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0

```

RESULT 11
US-09-949-016-150046
: Sequence 150046, Application US/09949016
: Patent No. 6812339
:
: GENERAL INFORMATION:
:
: APPLICANT: VENTER, J. Craig et al.
:
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
:
: FILE REFERENCE: CLO01307
:
: CURRENT APPLICATION NUMBER: US/09/949, 016
:
: CURRENT FILING DATE: 2000-04-14
:
: PRIOR APPLICATION NUMBER: 60/241,755
:
: PRIOR FILING DATE: 2000-10-20
:
: PRIOR APPLICATION NUMBER: 60/237,768
:

```



```

: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FASTSEQ for Windows Version 4.0.
: SEQ ID NO 150046
: LENGTH: 601
: TYPE: DNA
: ORGANISM: Human
: US-09-949-016-150046

```

Query Match	6.5%	Score 137,	DB 4;	Length 601;
Best Local Similarity	99.5%	Pred. NO. 6.3e-60;		
Matches 187;	Conservative	0;	Mismatches 1;	Indels 0;
				Gaps 0

QY	165	TTGAAAAGAGCTCAGACATTTCTTAAAGATGGATCTTAACTCATCTAAAGTTTC	182
Db	413	TTGAAAAGAGCTCAGACATTTCTTAAAGATGGATCTTAACTCATCTAAAGTTTC	472
QY	1825	TTCTCAAGAGATGCTCTGTTGGGAGAGGAGGCCCGACAAAGTATGTACAGACAAC	1884
Db	473	TTCTCAAGAGATGCTCTGTTGGGAGAGGAGGCCCGACAAAGTATGTGAGACAAC	532
QY	1885	ATCCAGCTTCAATGGCCAGCAGGTGGCGAAGATCTCTCCAGAGAGAAAGGCATATTTAT	1944
Db	533	ATCCAGCTTCAATGGCCAGCAGGTGGCGAAGATCTCTCCAGAGAGAAAGGCATATTTAT	592
QY	1945	GTGTGTGG 1952	
Db	593	GTGTGTGG 600	

```

US-09-949-016-150047
Sequence 150047 Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CLO01307
CURRENT APPLICATION NUMBER: US/09/949, 016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 150047
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-150047

```

Query Match	6.5%	Score 137	DB 4	Length 601
Best Local Similarity	99.5%	Pred. No. 6.3e-60		
Matches 187	Conservative 0	Mismatches 1	Indels 0	Gaps 0

Oy	1765	TTGGAAGAAGAGCTCAGACATTTCTTAAAGCATTTAAAGCTTAAAGGTTTC	182
Db	191	TTGGAAGAAGAGCTCAGACATTTCTTAAAGCATTTAAAGGTTTC	250
Oy	1825	TTCTCAAGAGATGCTCTGTTGGGAGAGGAAGCCCGCAAGATGTACAAGACAAC	188
Db	251	TTCTCAAGAGATGCTCTGTTGGGAGAGGAAGCCCGCAAGATGTACAAGACAAC	310
Oy	1885	ATCCAGCTTCATGCGCAGACAGGTGCGAGAAATCTCTCTCAGAGAAACGGCATTATTAAT	194
Db	311	ATCCAGCTTCATGCGCAGACAGGTGCGAGAAATCTCTCTCAGAGAAACGGCATTATTAAT	370

Qy	1945	GTGTGTGG	1952
Db	371	GTGTGTGG	378

RESULT 13  
US-09-949-016-150029

```

Sequence 150029, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949, 016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241, 755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237, 768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231, 498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 150029
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-150029

```

Query Match	6.0%	Score 125	DB 4	Length 601
Best Local Similarity	100.0%	Pred. No.	9.5e-54	
Matches 125	0	Mismatches	0	Gaps 0
Conservative	0	Indels	0	

Qy	779	AGAGAGAAAGCAAGATCTTGACATTCAGAGAGATCCAGTTTCAAGTGCATTTCAA	838
Db	379	AGAGAGAAAGCCAGATCTGTGACTTCAGCAGATCCAGTTTCAAGTGCATTTCAA	438
Qy	839	AGGCAATTGCACTTACTACGATGATGCGATPAAAAACCACTCTGCTGTGAATTGACA	898
Db	439	AGGCAAGTCACTTACTACGATGATGCGATPAAAAACCACTCTGCTGTGAATTGACA	498
Qy	899	TTTCA 903	
Db	499	TTTCA 503	

RESULT 14  
US-09-949-016-150041

```

: Sequence 150041: Application US/09949016
: Patent No. 6812339
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: CLO01307
: CURRENT APPLICATION NUMBER: US/09/949,016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 150041
: LENGTH: 601
: TYPE: DNA
: ORGANISM: Human
US-09-949-016-150041

```

Query Match 5.8%; Score 121; DB 4; Length 601;

Best Local Similarity 100.0%; Pred. No. 1.1e-51;  
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1556 AGATATCCATCTCTCTCGAACAACAATTTCTTCCACTTACGATGACCCCTCAATCC 1615

Db 124 AGATATCCATCTCTCTCGAACAACAATTTCTTCCACTTACGATGACCCCTCAATCC 183

QY 1616 CCATCATTAATGTTGGGTCGAGAAACCGGCATAGCCCGTTTATTGGGTTCTTCAACATA 1675

Db 184 CCATCATTAATGTTGGGTCGAGAAACCGGCATAGCCCGTTTATTGGGTTCTTCAACATA 243

QY 1676 G 1676

Db 244 G 244

## RESULT 15

US-09-949-016-150042

; Sequence 150042; Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: C0801307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 150042

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-150042

Query Match 5.8%; Score 121; DB 4; Length 601;

Best Local Similarity 100.0%; Pred. No. 1.1e-51;

Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1556 AGATATCCATCTCTCTCGAACAACAATTTCTTCCACTTACGATGACCCCTCAATCC 1615

Db 95 AGATATCCATCTCTCTCGAACAACAATTTCTTCCACTTACGATGACCCCTCAATCC 154

QY 1616 CCATCATTAATGTTGGGTCGAGAAACCGGCATAGCCCGTTTATTGGGTTCTTCAACATA 1675

Db 155 CCATCATTAATGTTGGGTCGAGAAACCGGCATAGCCCGTTTATTGGGTTCTTCAACATA 214

QY 1676 G 1676

Db 215 G 215

Search completed: August 27, 2005, 16:18:14  
Job time : 238.757 secs

1:	/cgn2_6/prodataa2/pubpna/us07_PUBCOMB.seq.*	Published Applications NA.*
2:	/cgn2_6/prodataa2/pubpna/PCF_NEW_PUB.seq.*	
3:	/cgn2_6/prodataa2/pubpna/US06_NEW_PUB.seq.*	
4:	/cgn2_6/prodataa2/pubpna/US06_PUBCOMB.seq.*	
5:	/cgn2_6/prodataa2/pubpna/US07_NEW_PUB.seq.*	
6:	/cgn2_6/prodataa2/pubpna/PCFUS_PUBCOMB.seq.*	
7:	/cgn2_6/prodataa2/pubpna/US08_NEW_PUB.seq.*	
8:	/cgn2_6/prodataa2/pubpna/US08_PUBCOMB.seq.*	
9:	/cgn2_6/prodataa2/pubpna/US09A_PUBCOMB.seq.*	
10:	/cgn2_6/prodataa2/pubpna/US09B_PUBCOMB.seq.*	
11:	/cgn2_6/prodataa2/pubpna/US09C_PUBCOMB.seq.*	
12:	/cgn2_6/prodataa2/pubpna/US09_NEW_PUB.seq.*	
13:	/cgn2_6/prodataa2/pubpna/US10A_PUBCOMB.seq.*	
14:	/cgn2_6/prodataa2/pubpna/US10B_PUBCOMB.seq.*	
15:	/cgn2_6/prodataa2/pubpna/US10D_PUBCOMB.seq.*	
16:	/cgn2_6/prodataa2/pubpna/US10F_PUBCOMB.seq.*	
17:	/cgn2_6/prodataa2/pubpna/US10F_PUBCOMB.seq.*	
18:	/cgn2_6/prodataa2/pubpna/US10F_PUBCOMB.seq.*	
19:	/cgn2_6/prodataa2/pubpna/US10G_PUBCOMB.seq.*	
20:	/cgn2_6/prodataa2/pubpna/US10H_PUBCOMB.seq.*	
21:	/cgn2_6/prodataa2/pubpna/US10I_PUBCOMB.seq.*	
22:	/cgn2_6/prodataa2/pubpna/US10I_NEW_PUB.seq.*	
23:	/cgn2_6/prodataa2/pubpna/US11A_PUBCOMB.seq.*	
24:	/cgn2_6/prodataa2/pubpna/US11_NEW_PUB.seq.*	
25:	/cgn2_6/prodataa2/pubpna/US60_NEW_PUB.seq.*	
26:	/cgn2_6/prodataa2/pubpna/US60_PUBCOMB.seq.*	

**Pred.** No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2097	100.0	2097	10	US-09-371-347-1	Sequence 1, Appl
2	2097	100.0	3259	10	US-09-371-347-24	Sequence 24, Appl
3	2097	100.0	2097	10	US-09-371-347-41	Sequence 41, Appl
4	2046	97.6	2097	10	US-09-371-347-43	Sequence 43, Appl
5	1854	88.4	2094	10	US-09-371-347-45	Sequence 45, Appl
6	1793	85.5	2093	10	US-09-371-347-47	Sequence 47, Appl
7	1062	50.6	3356	21	US-10-741-600-652	Sequence 692, Appl

8	1062	50.6	3274	21	US-10-741-600-693	Sequence 693, App
9	330	15.7	591	16	US-10-029-386-639	Sequence 639, App
10	328	15.6	379	16	US-10-029-386-20100	Sequence 20100, A
11	279	13.3	591	16	US-10-029-386-1735	Sequence 1735, Ap
12	277	13.2	379	16	US-10-029-386-15435	Sequence 15435, A
13	266	12.7	43985	21	US-10-741-600-17757	Sequence 17757, A
14	188	9.0	525	16	US-10-029-386-633	Sequence 633, App
15	175	8.3	175	16	US-10-029-386-14338	Sequence 14338, A
16	158	7.5	2475	10	US-09-909-567B-38	Sequence 38, Appl
17	158	7.5	2475	22	US-10-765-700-88	Sequence 88, Appl
18	158	7.5	21852	21	US-10-741-600-17986	Sequence 17986, A
19	150	7.2	201	21	US-10-741-600-15583	Sequence 15583, A
20	150	7.2	201	21	US-10-741-600-15584	Sequence 15584, A
21	150	7.2	201	21	US-10-741-600-15589	Sequence 15589, A
22	150	7.2	201	21	US-10-741-600-15590	Sequence 15590, A
23	150	7.2	201	21	US-10-741-600-15592	Sequence 15592, A
24	150	7.2	201	21	US-10-741-600-15593	Sequence 15593, A
25	150	7.2	201	21	US-10-741-600-15594	Sequence 15594, A
26	150	7.2	201	21	US-10-741-600-15598	Sequence 15598, A
27	150	7.2	201	21	US-10-741-600-15599	Sequence 15599, A
28	150	7.2	201	21	US-10-741-600-15600	Sequence 15600, A
29	150	7.2	201	21	US-10-741-600-15602	Sequence 15602, A
30	150	7.2	201	21	US-10-741-600-15606	Sequence 15606, A
31	150	7.2	201	21	US-10-741-600-15609	Sequence 15609, A
32	150	7.2	201	21	US-10-741-600-15610	Sequence 15610, A
33	150	7.2	201	21	US-10-741-600-15612	Sequence 15612, A
34	150	7.2	201	21	US-10-741-600-15613	Sequence 15613, A
35	150	7.2	201	21	US-10-741-600-15614	Sequence 15614, A
36	150	7.2	201	21	US-10-741-600-15620	Sequence 15620, A
37	150	7.2	201	21	US-10-741-600-15621	Sequence 15621, A
38	150	7.2	201	21	US-10-741-600-15623	Sequence 15623, A
39	150	7.2	201	21	US-10-741-600-15624	Sequence 15624, A
40	150	7.2	201	21	US-10-741-600-15625	Sequence 15625, A
41	150	7.2	201	21	US-10-741-600-15629	Sequence 15629, A
42	150	7.2	201	21	US-10-741-600-15630	Sequence 15630, A
43	150	7.2	201	21	US-10-741-600-15631	Sequence 15631, A
44	150	7.2	201	21	US-10-741-600-15633	Sequence 15633, A
45	150	7.2	201	21	US-10-741-600-15637	Sequence 15637, A

## ALIGNMENTS

```

1 RESULT 1
2 US-09-371-347-1
3 ; Sequence 1, Application US/09371347
4 ; Publication No. US20030082676A1
5 ; GENERAL INFORMATION:
6 ; APPLICANT: Roy A. Gravel et al.
7 ; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE REDUCTASE:
8 ; TITLE OF INVENTION: CLONING, AND METHODS FOR EVALUATING RISK OF NEURAL TUBE
9 ; TITLE OF INVENTION: DEFECTS, CARDIOVASCULAR DISEASE, AND CANCER
10 ; FILE REFERENCE: 5004/003003
11 ; CURRENT APPLICATION NUMBER: US/09/371,347
12 ; CURRENT FILING DATE: 1999-08-10
13 ; PRIOR APPLICATION NUMBER: 60/071,622
14 ; PRIOR FILING DATE: 1998-01-16
15 ; PRIOR APPLICATION NUMBER: 09/232,028
16 ; PRIOR FILING DATE: 1999-01-15
17 ; NUMBER OF SEQ ID NOS: 51
18 ; SOFTWARE: FastSeq for Windows Version 4.0
19 ; SEQ ID NO 1
20 ; LENGTH: 2097
21 ; TYPE: DNA
22 ; ORGANISM: Homo sapiens
23 US-09-371-347-1

```

	Query Match	100.0%	Score 2097	DB 10	Length 2097
	Beet Local Similarity	100.0%	Pred. No. 0		
	Matches 2097	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Q7	1	ATGAGGAGTTTCGTTACTATATATGCAACACAGCAGGACAGCAAGCAAGGCATCGCGAA	60		

Db	1	ATGAGGAGGTTCTTGTTACTATATGCTACACAGCAGGGAACAGGCAAGGCCATGCAGAA	60
Qy	61	GAATGTGTGAGCAAGCTGTGTACTATGATATTTCTGCAATCTTCACTGTATATGAA	120
Db	61	GAATGTGTGAGCAAGCTGTGTACTATGATATTTCTGCAATCTTCACTGTATATGAA	120
Qy	121	TCCGATATGATGACCTTAAACCGAAACAGCTCTCTGTGTGTGTGATTTTCAACAG	180
Db	121	TCCGATATGATGACCTTAAACCGAAACAGCTCTCTGTGTGTGTGATTTTCAACAG	180
Qy	181	GGCACCGGAGACCCACCCGACACAGCCCGCAAGTTTGTTAAGAAATACAAACAAACA	240
Db	181	GGCACCGGAGACCCACCCGACACAGCCCGCAAGTTTGTTAAGAAATACAAACAAACA	240
Qy	241	CTGCCGGTGAATTTCTTTGCTCACCTGGGGTATGGGTATCTGGGTCTCGGTATTCGAA	300
Db	241	CTGCCGGTGAATTTCTTTGCTCACCTGGGGTATGGGTATCTGGGTCTCGGTATTCGAA	300
Qy	301	TACACCTTCTTTTGCAATGGGGGGAGATATATTAAGACCTTCAAGAGCTTGGAGCC	360
Db	301	TACACCTTCTTTTGCAATGGGGGGAGATATATTAAGACCTTCAAGAGCTTGGAGCC	360
Qy	361	CGGCAATTTCTATGACACTGACATGCAAGATGACTGTGAGGTTTATGAACCTTGTGTGAG	420
Db	361	CGGCAATTTCTATGACACTGACATGCAAGATGACTGTGAGGTTTATGAACCTTGTGTGAG	420
Qy	421	CCGTGTGATTTGCTGACCTTGGCCAGCCCTCAAGAAACATTTTAGTCAAGCAGACAA	480
Db	421	CCGTGTGATTTGCTGACCTTGGCCAGCCCTCAAGAAACATTTTAGTCAAGCAGACAA	480
Qy	481	GAGGAGATTAATGGGGGCACTCCGGTGGCATCACCTGCATCTTGAGGACAGACCTGTG	540
Db	481	GAGGAGATTAATGGGGGCACTCCGGTGGCATCACCTGCATCTTGAGGACAGACCTGTG	540
Qy	541	AAGTCAAGAGCTGTACTACATTTGAATCTCAAGTCAAGCTTCTGAGATTCGATATTCAGA	600
Db	541	AAGTCAAGAGCTGTACTACATTTGAATCTCAAGTCAAGCTTCTGAGATTCGATATTCAGA	600
Qy	601	AGAAAGATTTCTGAGGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATCCATGTTGTA	660
Db	601	AGAAAGATTTCTGAGGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATCCATGTTGTA	660
Qy	661	ATTGAAGCTTGAAGTCTCTCACTTACCCGTTGAGTACCCCACTCTCAAGCCTCTGTG	720
Db	661	ATTGAAGCTTGAAGTCTCTCACTTACCCGTTGAGTACCCCACTCTCAAGCCTCTGTG	720
Qy	721	AATATTTCTGTTTACCCCCAGAAATATTTACAGTACATCTGCAGAGTCTCTTGGCCAG	780
Db	721	AATATTTCTGTTTACCCCCAGAAATATTTACAGTACATCTGCAGAGTCTCTTGGCCAG	780
Qy	781	GAGGAAGCCAGATCTGTGACTTCAGCAGATCCAGTTTTCAAGTCCCAATTTCAAG	840
Db	781	GAGGAAGCCAGATCTGTGACTTCAGCAGATCCAGTTTTCAAGTCCCAATTTCAAG	840
Qy	841	GCAATTCACCTTACTAGATATGATATGCAATTAACCACTCTGTGTGTATTTGACATTT	900
Db	841	GCAATTCACCTTACTAGATATGATATGCAATTAACCACTCTGTGTGTATTTGACATTT	900
Qy	901	TCAAATACAGACTTTTCTATACGCTGTGAGATGCTTCAAGGTGATCTGCCCTTACAGT	960
Db	901	TCAAATACAGACTTTTCTATACGCTGTGAGATGCTTCAAGGTGATCTGCCCTTACAGT	960
Qy	961	GATTTCTGAGATCAAAAGCTTACTCCAAAGACTGACGCTTGAAGATTAAGAGACACTGC	1020
Db	961	GATTTCTGAGATCAAAAGCTTACTCCAAAGACTGACGCTTGAAGATTAAGAGACACTGC	1020
Qy	1021	GTCTTTTGAATTAAGCAGACCAAAAGAAAGAGACTTATCCCAAGCATATA	1080
Db	1021	GTCTTTTGAATTAAGCAGACCAAAAGAAAGAGACTTATCCCAAGCATATA	1080
Qy	1081	CCTGGGGAGTGTCTCTCAGTTCATTTTACCTGTGTCTTGAATCCGAGCAATTCCT	1140
Db	1081	CCTGGGGAGTGTGTCTCTCAGTTCATTTTACCTGTGTCTTGAATCCGAGCAATTCCT	1140

QY	1144	AAAAAGCATTTTGGAGAGCCCTGAGGACATATACAGGTGACAGTGTCTGAAAACGGCAGG	1200
Dp	1141	AAAAAGCATTTTGGAGAGCCCTTGTGACATATACAGGTGACAGTGTCTGAAAACGGCAGG	1200
QY	1201	CTACAGAGCTGTGTGACATTAACAAGAGGGCAGCCGATTATAGCCGCTTTGTACGAGATGCC	1260
Dp	1201	CTACAGAGAGCTGTGTGACATTAACAAGAGGGCAGCCGATTATAGCCGCTTTGTACGAGATGCC	1260
QY	1261	TGTGCTGTCTTTGTGGATCTCTCTCTGCTTTCCCTTCTTGCCAGCCACCTAGTCTC	1320
Dp	1261	TGTGCTGTCTTTGTGGATCTCTCTCTCTGCTTTCCCTTCTTGCCAGCCACCTAGTCTC	1320
QY	1321	CTGTGTGAACATCTTCCCTAACTTCAACCCAGACCATATTCGTGTGCAAGCTCAAGTTTA	1380
Dp	1321	CTGTGTGAACATCTTCCCTAACTTCAACCCAGACCATATTCGTGTGCAAGCTCAAGTTTA	1380
QY	1381	TTTTCACCCAGAAAAGCTCCATTGTGTCTCAACATTTGTGTGAAATTTCTGTACTCTCCACA	1440
Dp	1381	TTTTCACCCAGAAAAGCTCCATTGTGTCTCAACATTTGTGTGAAATTTCTGTACTCTCCACA	1440
QY	1441	ACAGAGTTCTGCGGAAAGGAGATATGTACAGGCTGTGGCTGTGTGTGTGCTTCAAGTT	1500
Dp	1441	ACAGAGGTTCTGCGGAAAGGAGATATGTACAGGCTGTGGCTGTGTGTGTGCTTCAAGTT	1500
QY	1501	CTTACAGCAAAACATATACATGTACATCCCATGAAAGACAGCGGGAAAAGCCCTGGCTCTTAAGATA	1560
Dp	1501	CTTACAGCAAAACATATACATGTACATCCCATGAAAGACAGCGGGAAAAGCCCTGGCTCTTAAGATA	1560
QY	1561	TCCATCTCTCTCGACACACAAATTCCTTCCACTTACAGATGACCCCTCATCTCCATC	1620
Dp	1561	TCCATCTCTCTCGACACACAAATTCCTTCCACTTACAGATGACCCCTCATCTCCATC	1620
QY	1621	ATATGTGGGTCCAGGAAACCGGATGACCCCTTATTTGGGTTCCTAACACATATAGAG	1680
Dp	1621	ATATGTGGGTCCAGGAAACCGGATGACCCCTTATTTGGGTTCCTAACACATATAGAG	1680
QY	1681	AAACTCCAGAAACAACCCAGATGAGAAATTTTGGAGCAATGTGTTTGGCTGC	1740
Dp	1681	AAACTCCAGAAACAACCCAGATGAGAAATTTTGGAGCAATGTGTTTGGCTGC	1740
QY	1741	AGGATATAGATAGGAGTATATCTATTTGAGAAAAGCTCAGACATTTCTTAAAGCATGGG	1800
Dp	1741	AGGATATAGATAGGAGTATATCTATTTGAGAAAAGCTCAGACATTTCTTAAAGCATGGG	1800
QY	1801	ATCTTAACTCATCTAAAGTTCCTTCTCAAGAGATGCTCTGTGGGAGAGAGAGGCC	1860
Dp	1801	ATCTTAACTCATCTAAAGTTCCTTCTCAAGAGATGCTCTGTGGGAGAGAGAGGCC	1860
QY	1861	CCAGCAAAAGTATGTACAAACAACATCCAGTTTATGTGCGACAGGTGGCGAGAAATCCTC	1920
Dp	1861	CCAGCAAAAGTATGTACAAACAACATCCAGTTTATGTGCGACAGGTGGCGAGAAATCCTC	1920
QY	1921	CTCAGAGGAACGGCCATATTTATGTGTGTGAGATGCAAAAGAAATATGTGCCAAGATGTA	1980
Dp	1921	CTCAGAGGAACGGCCATATTTATGTGTGTGAGATGCAAAAGAAATATGTGCCAAGATGTA	1980
QY	1981	CATGATGCCCTTGTGCAAAATATATAGCAAGAAGGTTGGAGTTGAAAACTAGAAAGCAATG	2040
Dp	1981	CATGATGCCCTTGTGCAAAATATATAGCAAGAAGGTTGGAGTTGAAAACTAGAAAGCAATG	2040
QY	2041	AAAAACCCCTGGACATTTTAAAGAAAGAAAAGCTACCTTCACAGATATTTGGTCAATA	2097
Dp	2041	AAAAACCCCTGGACATTTTAAAGAAAGAAAAGCTACCTTCACAGATATTTGGTCAATA	2097

RESULT 2  
US-09-371-347-24  
; Sequence 24 Application US/09371347  
; Publication NO. US20030082676A1  
; GENERAL INFORMATION: Gravel et al.  
; APPLICANT: Roy A. Gravel et al.  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE REDUCTASE

TITLE OF INVENTION: CLONING, AND METHODS FOR EVALUATING RISK OF NEURAL TUBE  
TITLE OF INVENTION: DEFECTS, CARDIOVASCULAR DISEASE, AND CANCER  
FILE REFERENCE: 50004/003003  
CURRENT APPLICATION NUMBER: US/09/371,347  
CURRENT FILING DATE: 1999-08-10  
PRIOR APPLICATION NUMBER: 60/071,622  
PRIOR FILING DATE: 1998-01-16  
PRIOR APPLICATION NUMBER: 09/232,028  
PRIOR FILING DATE: 1999-01-15  
NUMBER OF SEQ ID NOS: 51  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 24  
LENGTH: 3259  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-371-347-24

Query Match 100.0%; Score 2097; DB 10; Length 3259;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGAGAGGTTCTGTACTATATGCTTACACAGCAGGAGCAGGCAAAAGCCATCGAGAA 60  
DB ATGAGAGAGGTTCTGTACTATATGCTTACACAGCAGGAGCAGGCAAAAGCCATCGAGAA 139  
QY 61 GAAATGTGAGAGAGCTGTGTACATGATTTTCTGCGAGATCTTCACTGATTAAGTAA 120  
DB 140 GAAATGTGAGAGAGCTGTGTACATGATTTTCTGCGAGATCTTCACTGATTAAGTAA 199  
QY 121 TCCGATATGATGACCTTAAACCCGAAACAGCTCTCTGTGTGTGTGTTTCTACACAG 180  
DB 200 TCCGATATGATGACCTTAAACCCGAAACAGCTCTCTGTGTGTGTGTTTCTACACAG 259  
QY 181 GGCAGCCGAGAGCCACCCGACACAGCCCGCAAGTTGTTAAGAAATACAGAACCAACA 240  
DB 260 GGCAGCCGAGAGCCACCCGACACAGCCCGCAAGTTGTTAAGAAATACAGAACCAACA 319  
QY 241 CTGCGCGGTGATTTCTTGTCTACCTGCGGTATGAGGTATCTGGGTCTCGGTATTCAGAA 300  
DB 320 CTGCGCGGTGATTTCTTGTCTACCTGCGGTATGAGGTATCTGGGTCTCGGTATTCAGAA 379  
QY 301 TACACCTACTTTGCAATGAGGAGGAGATTAATGATTAACGACTTCAAGAGCTTGAGCC 360  
DB 380 TACACCTACTTTGCAATGAGGAGGAGATTAATGATTAACGACTTCAAGAGCTTGAGCC 439  
QY 361 CGGCAATTTCTATGACACTGACATGAGATGACTGTGTAGTTTGAACCTTGTGTGAG 420  
DB 440 CGGCAATTTCTATGACACTGACATGAGATGACTGTGTAGTTTGAACCTTGTGTGAG 499  
QY 421 CCGTGTATGCTGAGACTCTGCGCAGCCCTCAAGAAAGCATTTTGTGTCAGAGAGACAA 480  
DB 500 CCGTGTATGCTGAGACTCTGCGCAGCCCTCAAGAAAGCATTTTGTGTCAGAGAGACAA 559  
QY 481 GAGAGATAGTATGAGGAGCACTCCGAGTGCATCACTGCACTTGAAGGACAGACTTGTG 540  
DB 560 GAGAGATAGTATGAGGAGCACTCCGAGTGCATCACTGCACTTGAAGGAGAGACTTGTG 619  
QY 541 AAGTCAGAGCTGTACATGATGATCTCAAGTCAAGCTTCTGAGATTCGATGATTCAGGA 600  
DB 620 AAGTCAGAGCTGTACATGATGATCTCAAGTCAAGCTTCTGAGATTCGATGATTCAGGA 679  
QY 601 AGAAAGATTTCTGAGGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATCGAATGTTGA 660  
DB 680 AGAAAGATTTCTGAGGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATCGAATGTTGA 739  
QY 661 ATGAGAGCTTTGAGTCTCACTTACCGGTCCGATACCCCACTCTCAAGAGCTCTCTG 720  
DB 740 ATGAGAGCTTTGAGTCTCACTTACCGGTCCGATACCCCACTCTCAAGAGCTCTCTG 799  
QY 721 AATATTCCTGGTTTACCCCAAGATATTTAAGGTATCATCTGAGAGAGTCTCTTGAGCAG 780  
DB 800 AATATTCCTGGTTTACCCCAAGATATTTAAGGTATCATCTGAGAGAGTCTCTTGAGCAG 859

QY 781 GAGAAAGCCAGATATCTGTGACTTCAAGAGATCCAGTTTCAAGTGCATTTCAAG 840  
DB 860 GAGAAAGCCAGATATCTGTGACTTCAAGAGATCCAGTTTCAAGTGCATTTCAAG 919  
QY 841 GCAATTCATTAATGAGATATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
DB 920 GCAATTCATTAATGAGATATGATGATGATGATGATGATGATGATGATGATGATGAT 979  
QY 901 TGAATATGAGATTTTCTGATGAGGATGATGATGATGATGATGATGATGATGATGAT 960  
DB 980 TGAATATGAGATTTTCTGATGAGGATGATGATGATGATGATGATGATGATGATGAT 1039  
QY 961 GATTCAGATTAAGAGCTTCAAGAGATGATGATGATGATGATGATGATGATGATGATG 1020  
DB 1040 GATTCAGATTAAGAGCTTCAAGAGATGATGATGATGATGATGATGATGATGATGATG 1099  
QY 1021 GTCTTTTGAATTAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080  
DB 1100 GTCTTTTGAATTAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1159  
QY 1081 CCGCGGAGATGATCTCTGAGTTCATTTTACCTGATCTGAAATCCGAGCATTCCT 1140  
DB 1160 CCGCGGAGATGATCTCTGAGTTCATTTTACCTGATCTGAAATCCGAGCATTCCT 1219  
QY 1141 AAAAAGCATTTTGTGAGAGCCCTTGTGAGATATACAGAGAGAGAGAGAGAGAGAG 1200  
DB 1220 AAAAAGCATTTTGTGAGAGCCCTTGTGAGATATACAGAGAGAGAGAGAGAGAGAG 1279  
QY 1201 CTACAGAGCTGTGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260  
DB 1280 CTACAGAGCTGTGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1339  
QY 1261 TGTGCTGCTGTGTGAGATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320  
DB 1340 TGTGCTGCTGTGTGAGATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1399  
QY 1321 CTGCTGAGATCTTCTTAACTTCAACCCAGAGATATGCTGAGAGAGAGAGAGAG 1380  
DB 1400 CTGCTGAGATCTTCTTAACTTCAACCCAGAGATATGCTGAGAGAGAGAGAGAG 1459  
QY 1381 TTTACCCAGAGAGAGCTTCAATTTTGTCTTCAATTTGAGAGATTTCTGATCTGAC 1440  
DB 1460 TTTACCCAGAGAGAGCTTCAATTTTGTCTTCAATTTGAGAGATTTCTGATCTGAC 1519  
QY 1441 ACAGAGCTTCTGAG 1500  
DB 1520 ACAGAGCTTCTGAG 1579  
QY 1501 CTTCAGCCAAACATGATGATCCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560  
DB 1580 CTTCAGCCAAACATGATGATCCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1639  
QY 1561 TCCATCTCTCTGAG 1620  
DB 1640 TCCATCTCTCTGAG 1699  
QY 1621 ATATATGAGGATGAG 1680  
DB 1700 ATATATGAGGATGAG 1759  
QY 1681 AAACTCAG 1740  
DB 1760 AAACTCAG 1819  
QY 1741 AGGATTAAGATTAAGAGATTAATCTTCAAGAAAGAGAGAGAGAGAGAGAGAGAG 1800  
DB 1820 AGGATTAAGATTAAGAGATTAATCTTCAAGAAAGAGAGAGAGAGAGAGAGAGAG 1879  
QY 1801 ATCTTAATCATTAAGAGATTTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860  
DB 1880 ATCTTAATCATTAAGAGATTTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1939  
QY 1861 CAGAGAGATATGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920

Db 1940 CCGAGCAAGATGATGACAAAGCAACATCAGCTTCATGCGACAGAGGTGGAGAAATCTTC 1999  
Qy 1921 CTCACAGAGAACGGCCATATTTATGTGTGAGATGCAAAAGATATGSCCAAGATGTA 1980  
Db 2000 CTCACAGAGAACGGCCATATTTATGTGTGAGATGCAAAAGATATGSCCAAGATGTA 2059  
Qy 1981 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGGTTGAGTTGAAAACTAGAAGCATG 2040  
Db 2060 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGGTTGAGTTGAAAACTAGAAGCATG 2119  
Qy 2041 AAAACCTGGCCACTTTAAAAAGAAAGAAAGCGTACCTTCAGAGTATTTGTCATTA 2097  
Db 2120 AAAACCTGGCCACTTTAAAAAGAAAGAAAGCGTACCTTCAGAGTATTTGTCATTA 2176

RESULT 3  
US-09-371-347-41  
; Sequence 41, Application US/09371347  
; Publication No. US2003082676A1  
; GENERAL INFORMATION:  
; APPLICANT: Roy A. Gravel et al.  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE REDUCTASE:  
; TITLE OF INVENTION: CLONING, AND METHODS FOR EVALUATING RISK OF NEURAL TUBE  
; TITLE OF INVENTION: DEFECTS, CARDIOVASCULAR DISEASE, AND CANCER  
; FILE REFERENCE: 50004/003003  
; CURRENT APPLICATION NUMBER: US/09/371,347  
; CURRENT FILING DATE: 1999-08-10  
; PRIOR APPLICATION NUMBER: 60/071,622  
; PRIOR FILING DATE: 1998-01-16  
; PRIOR APPLICATION NUMBER: 09/232,028  
; PRIOR FILING DATE: 1999-01-15  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: FaSeq for Windows Version 4.0  
; SEQ ID NO 41  
; LENGTH: 2097  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-09-371-347-41

Query Match 97.6%; Score 2046; DB 10; Length 2097;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2096; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGAGAGGTTCTGTACTATATATGCTACACAGCAGGACACAGCAAGCCATGCGAGAA 60  
Db 1 ATGAGAGGTTCTGTACTATATATGCTACACAGCAGGACACAGCAAGCCATGCGAGAA 60  
Qy 61 GAATATGTGAGCAAGCTGTGTGATCATGAGATTTCTGCAATCTTCACTGTATTAAGAA 120  
Db 61 GAATATGTGAGCAAGCTGTGTGATCATGAGATTTCTGCAATCTTCACTGTATTAAGAA 120  
Qy 121 TCGCATTAAGTATGACTTAAAAACCGAAACAGCTCTCTGTGTGTGTGTTGTTCTACACAG 180  
Db 121 TCGCATTAAGTATGACTTAAAAACCGAAACAGCTCTCTGTGTGTGTGTTGTTCTACACAG 180  
Qy 181 GGCACCGGAGACCCACCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACA 240  
Db 181 GGCACCGGAGACCCACCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACA 240  
Qy 241 CTGCGCGGTGATTTCTTTGCTCACCTGCGGTATAGGTTTCTGGGCTCTCGGTGATTCAGAA 300  
Db 241 CTGCGCGGTGATTTCTTTGCTCACCTGCGGTATAGGTTTCTGGGCTCTCGGTGATTCAGAA 300  
Qy 301 TACACTACTTTTTCAGATGGGGGAGAGATTAATGATTAAGACTTCAAGAGCTTGAAGCC 360  
Db 301 TACACTACTTTTTCAGATGGGGGAGAGATTAATGATTAAGACTTCAAGAGCTTGAAGCC 360  
Qy 361 CGGCAATTTCTATGACATGACATGACATGCTGTAGTGTAAAGCTTGTGGTTGAG 420  
Db 361 CGGCAATTTCTATGACATGACATGACATGCTGTAGTGTAAAGCTTGTGGTTGAG 420  
Qy 421 CCGTGAATTTGCTGGAATCTGTGCGCAGCCCTCAGAAAGCAATTTAGGTCAAGCAGAGCAA 480

Db 421 CCGTGAATTTGCTGGAATCTGTGCGCAGCCCTCAGAAAGCAATTTAGGTCAAGCAGAGCAA 480  
Qy 481 GAGAGATTAAGTGGCCATCTCCCGGTGAGCATCACTGCATCTTGTAGAGACAGACTTGTG 540  
Db 481 GAGAGATTAAGTGGCCATCTCCCGGTGAGCATCACTGCATCTTGTAGAGACAGACTTGTG 540  
Qy 541 AAGTCAGAGCTGTACATCAATTAATTCAGATGAGCTTCTGAGATTCGATGATTCAGGA 600  
Db 541 AAGTCAGAGCTGTACATCAATTAATTCAGATGAGCTTCTGAGATTCGATGATTCAGGA 600  
Qy 601 AGAAGAGTTCTGAGGTTTGAAGCAAAATGCAAGTGAAGCAAGCAACCAATGATGTTGA 660  
Db 601 AGAAGAGTTCTGAGGTTTGAAGCAAAATGCAAGTGAAGCAAGCAACCAATGATGTTGA 660  
Qy 661 ATTGAAGCTTTGAGTCTCATCTTACCCGTTGAGTACCCCACTCTCAAGAGCTCTCTG 720  
Db 661 ATTGAAGCTTTGAGTCTCATCTTACCCGTTGAGTACCCCACTCTCAAGAGCTCTCTG 720  
Qy 721 AATATTCCTGTTTACCCCAAGATATTTTACAGTACATCTGACAGAGTCTCTGGCCAG 780  
Db 721 AATATTCCTGTTTACCCCAAGATATTTTACAGTACATCTGACAGAGTCTCTGGCCAG 780  
Qy 781 GAGAAAGCCAGTATCTGTGACTTCAGAGATCCAGTTTCAAGTGCATTTGAAG 840  
Db 781 GAGAAAGCCAGTATCTGTGACTTCAGAGATCCAGTTTCAAGTGCATTTGAAG 840  
Qy 841 GCAGTTCAACTTACTACGATGATGCAATTAACCACTCTGCTGTAGATTTGACATT 900  
Db 841 GCAGTTCAACTTACTACGATGATGCAATTAACCACTCTGCTGTAGATTTGACATT 900  
Qy 901 TCAATATCAGACTTTTCTATCAGCTGTGAGATGCTTCAAGCTGATCTGCCATACAGT 960  
Db 901 TCAATATCAGACTTTTCTATCAGCTGTGAGATGCTTCAAGCTGATCTGCCATACAGT 960  
Qy 961 GATTCGAGGTCAAGAGCTCTCCCAAGATGCGAGCTTGAAGATTAAGAGAGCACTGC 1020  
Db 961 GATTCGAGGTCAAGAGCTCTCCCAAGATGCGAGCTTGAAGATTAAGAGAGCACTGC 1020  
Qy 1021 GTCCCTTTGAAATTAAGGACAGACCAAGAAAGAGAGCTTACCTCCAGCATATA 1080  
Db 1021 GTCCCTTTGAAATTAAGGACAGACCAAGAAAGAGAGCTTACCTCCAGCATATA 1080  
Qy 1081 CCGCGGAGTGTCTCTCAAGTCAATTTTCTCTGCTGTGAAATCCGAGCAATTCCT 1140  
Db 1081 CCGCGGAGTGTCTCTCAAGTCAATTTTCTCTGCTGTGAAATCCGAGCAATTCCT 1140  
Qy 1141 AAAAAGCAATTTTGGAGGCCCTGTGGAATATACAGTGAAGTGTGAAAAGCGCAGG 1200  
Db 1141 AAAAAGCAATTTTGGAGGCCCTGTGGAATATACAGTGAAGTGTGAAAAGCGCAGG 1200  
Qy 1201 CTACAGAGCTGTGACATTAACAGAGGAGCAGCCGATTAATAGCCCTTGTATGAGATGCC 1260  
Db 1201 CTACAGAGCTGTGACATTAACAGAGGAGCAGCCGATTAATAGCCCTTGTATGAGATGCC 1260  
Qy 1261 TGTGCTGCTGTGTGATCTCTCTGCTTCTCTTCCCTTCCAGGACCACTCAAGTCTC 1320  
Db 1261 TGTGCTGCTGTGTGATCTCTCTGCTTCTCTTCCCTTCCAGGACCACTCAAGTCTC 1320  
Qy 1321 CTGCTGAACATCTTCTTAACTTCAACCCAGACCAATTCGTGTGACAGCTCAAGTTTA 1380  
Db 1321 CTGCTGAACATCTTCTTAACTTCAACCCAGACCAATTCGTGTGACAGCTCAAGTTTA 1380  
Qy 1381 TTTCAACCCAGAAAGCTCAATTTTGTCTTCAACATTTGGAATTTCTGTCTACATGCGACA 1440  
Db 1381 TTTCAACCCAGAAAGCTCAATTTTGTCTTCAACATTTGGAATTTCTGTCTACATGCGACA 1440  
Qy 1441 ACAGAGTCTCTGGAAGGAGATGATGATACAGGCTGAGCTGTGTGTTGCTTCAAGTT 1500  
Db 1441 ACAGAGTCTCTGGAAGGAGATGATGATGATACAGGCTGAGCTGTGTGTTGCTTCAAGTT 1500  
Qy 1501 CTTCAAGCAACATTAATGATGATCCCATGAAGACAGGGGAAAACCTGTGCTCTTAAGATA 1560

Db 1501 CTTGAGCCAAAGATACATCATCCATGAAGACAGCGGAAAGCCCTGCTCTTAAGATA 1560  
Qy 1561 TCCATCTCTCTGAGCAACAATTCCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
Db 1561 TCCATCTCTCTGAGCAACAATTCCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
Qy 1621 ATAAATGTTGGTTCAGGAACCGGATAGCCCGTTATTTGGTTCCTAACAATAGAG 1680  
Db 1621 ATAAATGTTGGTTCAGGAACCGGATAGCCCGTTATTTGGTTCCTAACAATAGAG 1680  
Qy 1681 AAATCTCAAGAACACACCCAGATGGAATTTTGAAGCAATGTGTTTTTGGCTGC 1740  
Db 1681 AAATCTCAAGAACACACCCAGATGGAATTTTGAAGCAATGTGTTTTTGGCTGC 1740  
Qy 1741 AAGCATTAAGATAGGATTAATCTATTCAGAAAAAGACTGACATTTCTTAAAGATGAG 1800  
Db 1741 AAGCATTAAGATAGGATTAATCTATTCAGAAAAAGACTGACATTTCTTAAAGATGAG 1800  
Qy 1801 ATCTTAATCATCTAAAGGTTTCTTCTGAGAGATGCTCTGTTGGGAGAGAGAGCC 1860  
Db 1801 ATCTTAATCATCTAAAGGTTTCTTCTGAGAGATGCTCTGTTGGGAGAGAGAGCC 1860  
Qy 1861 CCAGCAAAAGTATGTACAGAACATTCAGCTTCATGCGCAGAGGTGGCGAATCTTC 1920  
Db 1861 CCAGCAAAAGTATGTACAGAACATTCAGCTTCATGCGCAGAGGTGGCGAATCTTC 1920  
Qy 1921 CTCGAGAGAACGGCCATTTATTTGTGTGTGAGATGCAAAAGATTTGGCCAAAGATGTA 1980  
Db 1921 CTCGAGAGAACGGCCATTTATTTGTGTGTGAGATGCAAAAGATTTGGCCAAAGATGTA 1980  
Qy 1981 CATGATGCTTGTGCAAAATTAATTAAGCAAAAGGTGAGTTGAAAACTAGAAAGATG 2040  
Db 1981 CATGATGCTTGTGCAAAATTAATTAAGCAAAAGGTGAGTTGAAAACTAGAAAGATG 2040  
Qy 2041 AAAACCTGGCCACTTTAAAGAAAGAAAAAGCTTACCTTCAAGATTTTGTGATTA 2097  
Db 2041 AAAACCTGGCCACTTTAAAGAAAGAAAAAGCTTACCTTCAAGATTTTGTGATTA 2097

RESULT 4  
US-09-371-347-43  
; Sequence 43, Application US/09371347  
; Publication No. US20030082676A1  
; GENERAL INFORMATION:  
; APPLICANT: Roy A. Gravel et al.  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE REDUCTASE:  
; TITLE OF INVENTION: CLONING, AND METHODS FOR EVALUATING RISK OF NEURAL TUBE  
; FILE REFERENCE: 50004/003003  
; CURRENT APPLICATION NUMBER: US/09/371,347  
; PRIOR FILING DATE: 1999-08-10  
; PRIOR APPLICATION NUMBER: 60/071,622  
; PRIOR FILING DATE: 1998-01-16  
; PRIOR APPLICATION NUMBER: 09/232,028  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 43  
; LENGTH: 2097  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-371-347-43

Query Match 97.6%; Score 2046; DB 10; Length 2097;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2096; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATAGAGAGGTTTCTGTAATATATGCTACACAGAGGACAGCAAGGCATGCGAGAA 60  
Db 1 ATAGAGAGGTTTCTGTAATATATGCTACACAGAGGACAGCAAGGCATGCGAGAA 60  
Qy 61 GAATGTGTAGCAAGCTGTGTATGATGATTTTCTGAGATCTTCACTGATTAAGTAA 120  
Db 61 GAATGTGTAGCAAGCTGTGTATGATGATTTTCTGAGATCTTCACTGATTAAGTAA 120

Db 61 GAATGTGTAGCAAGCTGTGTATGATGATTTTCTGAGATCTTCACTGATTAAGTAA 120  
Qy 121 TCCGATTAAGTATGACCTAAAAACGAAACAGCTCCCTTGTGTTGTGTTTCTACAG 180  
Db 121 TCCGATTAAGTATGACCTAAAAACGAAACAGCTCCCTTGTGTTGTGTTTCTACAG 180  
Qy 181 GGCACCGGAGACCCACCGGACACAGCCCGCAAGTTTGTTAAGAAATACAGAACCAACA 240  
Db 181 GGCACCGGAGACCCACCGGACACAGCCCGCAAGTTTGTTAAGAAATACAGAACCAACA 240  
Qy 241 CTGCGGTTGATTTCTTGTCTCACTCGGATATGAGTTTCTGAGTTTCAAG 300  
Db 241 CTGCGGTTGATTTCTTGTCTCACTCGGATATGAGTTTCTGAGTTTCAAG 300  
Qy 301 TACACCTACTTTTGAATGGGGGAAAGATTAATGATTAACGCTTCAAGAGCTGGAGCC 360  
Db 301 TACACCTACTTTTGAATGGGGGAAAGATTAATGATTAACGCTTCAAGAGCTGGAGCC 360  
Qy 361 CGGCAATTTCTATGACATGACATGACATGACATGACATGACATGACATGACATGAC 420  
Db 361 CGGCAATTTCTATGACATGACATGACATGACATGACATGACATGACATGACATGAC 420  
Qy 421 CCGTGATTTGCTGACTCTGCGCAGCCCTCAGAAAGCATTTTAAAGTCAAGAGAGACA 480  
Db 421 CCGTGATTTGCTGACTCTGCGCAGCCCTCAGAAAGCATTTTAAAGTCAAGAGAGACA 480  
Qy 481 GAGGATTAAGTGGGCACTCCCGGTGGATACCTGATCCTTGAAGACACACCTGTG 540  
Db 481 GAGGATTAAGTGGGCACTCCCGGTGGATACCTGATCCTTGAAGACACACCTGTG 540  
Qy 541 AAGTCAGAGCTCTACATTAATCTCAATGAGCTTCTGAGATTCAGATTCAGAGA 600  
Db 541 AAGTCAGAGCTCTACATTAATCTCAATGAGCTTCTGAGATTCAGATTCAGAGA 600  
Qy 541 AAGTCAGAGCTCTACATTAATCTCAATGAGCTTCTGAGATTCAGATTCAGAGA 600  
Db 541 AAGTCAGAGCTCTACATTAATCTCAATGAGCTTCTGAGATTCAGATTCAGAGA 600  
Qy 601 AAAAAAGATTCGAGGTTTGAAGCAAAATGCAATGACAGCAACCAATTCATTTGTA 660  
Db 601 AAAAAAGATTCGAGGTTTGAAGCAAAATGCAATGACAGCAACCAATTCATTTGTA 660  
Qy 661 ATTGAAGACTTTGAGTCTCTCACTTACCGGTGGATACCCCACTCTCAAGCCTCTCG 720  
Db 661 ATTGAAGACTTTGAGTCTCTCACTTACCGGTGGATACCCCACTCTCAAGCCTCTCG 720  
Qy 721 AATATTCGTTTACCCCGCAGATATTTAAGATGATCTGAGAGTCTTGGCCAG 780  
Db 721 AATATTCGTTTACCCCGCAGATATTTAAGATGATCTGAGAGTCTTGGCCAG 780  
Qy 781 GAGGAAAGCAATCTGTGACTTCAAGATTCAGATTCAGATTCAGATTCAGATTCAG 840  
Db 781 GAGGAAAGCAATCTGTGACTTCAAGATTCAGATTCAGATTCAGATTCAGATTCAG 840  
Qy 841 GCGATTCATCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
Db 841 GCGATTCATCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
Qy 901 TCAAAATACAGATTTTCTCATGAGCTGAGATGCTTCAAGGATTCAGATTCAGAT 960  
Db 901 TCAAAATACAGATTTTCTCATGAGCTGAGATGCTTCAAGGATTCAGATTCAGAT 960  
Qy 961 GATTCGAGTACCAAGCTTCTCAAGATTCAGATTCAGATTCAGATTCAGATTCAG 1020  
Db 961 GATTCGAGTACCAAGCTTCTCAAGATTCAGATTCAGATTCAGATTCAGATTCAG 1020  
Qy 1021 GTTCCTTTTAAATTAAGGACACACAAAGAAAGAGATTCCTTACCCAGCATTA 1080  
Db 1021 GTTCCTTTTAAATTAAGGACACACAAAGAAAGAGATTCCTTACCCAGCATTA 1080  
Qy 1081 CCGCGGAGATTTCTCTCAAGTTCATTTTACCTGATTCGATTCGATTCGATTCGAT 1140  
Db 1081 CCGCGGAGATTTCTCTCAAGTTCATTTTACCTGATTCGATTCGATTCGATTCGAT 1140  
Qy 1141 AAAAAGCAATTTTGAAGCCCTTGTGACATTAACAGTACAGTGTGAAAAAGCGCAG 1200  
Db 1141 AAAAAGCAATTTTGAAGCCCTTGTGACATTAACAGTACAGTGTGAAAAAGCGCAG 1200



QY 1201 CTACAGAGCTGTGCACTAAACAAGGGGAGCCGATATAGCCGCTTTGTACAGATGCC 1260  
DB 1201 CTACAGAGCTGTGCACTAAACAAGGGGAGCCGATATAGCCGCTTTGTACAGATGCC 1260  
QY 1261 TGTGCTGTGCTTTGTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1320  
DB 1261 TGTGCTGTGCTTTGTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1320  
QY 1321 CTGCTGACATCTTCTTAACTTCAACCCAGACATATTCGTGTGACGCTCAAGTTTA 1380  
DB 1321 CTGCTGACATCTTCTTAACTTCAACCCAGACATATTCGTGTGACGCTCAAGTTTA 1380  
QY 1381 TTTTACCAGAGAAAGCTCCATTTTGTCTTCAACATTTGTGAAATTTCTGTCTACGCCA 1440  
DB 1381 TTTTACCAGAGAAAGCTCCATTTTGTCTTCAACATTTGTGAAATTTCTGTCTACGCCA 1440  
QY 1441 ACAGAGTTTCTGCGAAGGAGATGTACAGGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1500  
DB 1441 ACAGAGTTTCTGCGAAGGAGATGTACAGGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1500  
QY 1501 CTTCAGCCAAACATACATGATCCCATGAAAGAGCGGAAAGCCCTGCTCTTCAAGATA 1560  
DB 1501 CTTCAGCCAAACATACATGATCCCATGAAAGAGCGGAAAGCCCTGCTCTTCAAGATA 1560  
QY 1561 TCCATCTCTCTCGAACAAACAAATCTTTTCACTTACAGATGACCCCTCAATCCCATC 1620  
DB 1561 TCCATCTCTCTCGAACAAACAAATCTTTTCACTTACAGATGACCCCTCAATCCCATC 1620  
QY 1621 ATATATGTGGGTCCAGAGACCGGATAGCCCGCTTTTGTGGGTCTTCAACATAGAG 1680  
DB 1621 ATATATGTGGGTCCAGAGACCGGATAGCCCGCTTTTGTGGGTCTTCAACATAGAG 1680  
QY 1681 AAATCTCAAGAACAAACCCAGATGAAATTTTGAAGCAATGTGTGTTTGTGTGCTGC 1740  
DB 1681 AAATCTCAAGAACAAACCCAGATGAAATTTTGAAGCAATGTGTGTTTGTGTGCTGC 1740  
QY 1741 AGGATTAAGATAGGATTAATCTATTCAGAAAGAGCTCAGACATTTCTTTAAGCATGG 1800  
DB 1741 AGGATTAAGATAGGATTAATCTATTCAGAAAGAGCTCAGACATTTCTTTAAGCATGG 1800  
QY 1801 ATCTTAATCTATCTTAAGGTTTCTTCTCAAGATATCTCTGTGGGAGAGAGAGGCC 1860  
DB 1801 ATCTTAATCTATCTTAAGGTTTCTTCTCAAGATATCTCTGTGGGAGAGAGAGGCC 1860  
QY 1861 CCAGCAAGATATGACAAAGACATCAGCTTATGAGCCAGAGAGAGAGAGATCTTC 1920  
DB 1861 CCAGCAAGATATGACAAAGACATCAGCTTATGAGCCAGAGAGAGAGAGATCTTC 1920  
QY 1921 CTCCAGAGAAAGGCAATATTTATGTGTGAGATGCAAGAAATATGCGCAAGATGTA 1980  
DB 1921 CTCCAGAGAAAGGCAATATTTATGTGTGAGATGCAAGAAATATGCGCAAGATGTA 1980  
QY 1981 CATATATCTCTTGTGCAATTAATAGCAAGAGGTGTGAGTGAATACTAAGAGCATG 2040  
DB 1981 CATATATCTCTTGTGCAATTAATAGCAAGAGGTGTGAGTGAATACTAAGAGCATG 2040  
QY 2041 AAAACCTGGCCCTTTAAAGAGAAAGAGCTTACCTTCAGATTTTGTGCTATA 2097  
DB 2041 AAAACCTGGCCCTTTAAAGAGAAAGAGCTTACCTTCAGATTTTGTGCTATA 2097

## RESULT 5

US-09-371-347-45  
; Sequence 45, Application US/09371347  
; Publication No. US20030082676A1  
; GENERAL INFORMATION:  
; APPLICANT: Roy A. Gravel et al.  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE REDUCTASE:  
; TITLE OF INVENTION: CLONING, AND METHODS FOR EVALUATING RISK OF NEURAL TUBE  
; TITLE OF INVENTION: DEFECTS, CARDIOVASCULAR DISEASE, AND CANCER  
; FILE REFERENCE: 50004/003003  
; CURRENT APPLICATION NUMBER: US/09/371.347

; CURRENT FILING DATE: 1999-08-10  
; PRIOR APPLICATION NUMBER: 60/071,622  
; PRIOR FILING DATE: 1998-01-16  
; PRIOR APPLICATION NUMBER: 09/232,028  
; PRIOR FILING DATE: 1999-01-15  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 45  
; LENGTH: 2094  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-371-347-45

Query Match 88.4%; Score 1854; DB 10; Length 2094;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2094; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 ATGAGAGGTTTCTGTACTATATGCTTACACAGAGGAGCAGGCAAGGCCATCGAGAA 60  
DB 1 ATGAGAGGTTTCTGTACTATATGCTTACACAGAGGAGCAGGCAAGGCCATCGAGAA 60  
QY 61 GAAATGTGAGCAGCTGTGATGATGATTTTCTGAGATCTTCACTGTATTAGTAA 120  
DB 61 GAAATGTGAGCAGCTGTGATGATGATTTTCTGAGATCTTCACTGTATTAGTAA 120  
QY 121 TCCGATTAATGATGACCTTAAACCCGAAACAGCTCTTGTGTGTGTGTTTCTACACG 180  
DB 121 TCCGATTAATGATGACCTTAAACCCGAAACAGCTCTTGTGTGTGTGTTTCTACACG 180  
QY 181 GGGACCGGAGACCCACCCGACACAGCCCGCAAGTTTGTAAAGAAATACAAACCA 240  
DB 181 GGGACCGGAGACCCACCCGACACAGCCCGCAAGTTTGTAAAGAAATACAAACCA 240  
QY 241 CTGCGGTGATTTCTTGTCTCACTGCGGTATGAGTTTCTGAGTTTCTGAGTTTCA 300  
DB 241 CTGCGGTGATTTCTTGTCTCACTGCGGTATGAGTTTCTGAGTTTCTGAGTTTCA 300  
QY 301 TACACCTACTTTTGAATGAGGAGGAGATTAATTAAGATTAAGATTAAGATTAAG 360  
DB 301 TACACCTACTTTTGAATGAGGAGGAGATTAATTAAGATTAAGATTAAGATTAAG 360  
QY 361 CGGATTTCTATGACCTGACATGAGATGATGATGATGATGATGATGATGATGATG 420  
DB 361 CGGATTTCTATGACCTGACATGAGATGATGATGATGATGATGATGATGATGATG 420  
QY 421 CCGTGAATGCTGATCTGAGCCAGCCCTCAGAAACATTTTGAAGTCAAGCAGAGCAA 480  
DB 421 CCGTGAATGCTGATCTGAGCCAGCCCTCAGAAACATTTTGAAGTCAAGCAGAGCAA 480  
QY 481 GAGAGATTAAGTGGGCACTCCGCTGAGATCACTGATCTTGAAGACAGACTTTGTG 540  
DB 481 GAGAGATTAAGTGGGCACTCCGCTGAGATCACTGATCTTGAAGACAGACTTTGTG 540  
QY 541 AATCTGAGTGTGTAACATTAATGATCTCAAGTGAATCTTGAAGTGAATGATGATG 600  
DB 541 AATCTGAGTGTGTAACATTAATGATCTCAAGTGAATCTTGAAGTGAATGATGATG 600  
QY 601 AGAAGGATCTGAGGTTTGAAGCAAAATGAGTGAACAGCAACCAATCAATGTTGTA 660  
DB 601 AGAAGGATCTGAGGTTTGAAGCAAAATGAGTGAACAGCAACCAATCAATGTTGTA 660  
QY 661 ATTGAAGATTTGAGTCTTCACTTACCCGTTGAGTACCCCACTCTCAAGCTCTCTG 720  
DB 661 ATTGAAGATTTGAGTCTTCACTTACCCGTTGAGTACCCCACTCTCAAGCTCTCTG 720  
QY 721 AATATTCGTGTTTACCCCGCAATTTTACAGATATCTGAGAGAGTCTTGTGGCAG 780  
DB 721 AATATTCGTGTTTACCCCGCAATTTTACAGATATCTGAGAGAGTCTTGTGGCAG 780  
QY 781 GAGGAAAGCAATATCTGATCTTCAAGCAGATCAAGTTTCAAGTCAAGTTTCAAG 840  
DB 781 GAGGAAAGCAATATCTGATCTTCAAGCAGATCAAGTTTCAAGTCAAGTTTCAAG 840

QY 841 GGAGTTCAACTTAAGATGATGCGATTAACCACTCTGCTGATGATGACAT 900  
DB 841 GGAGTTCAACTTAAGATGATGCGATTAACCACTCTGCTGATGATGACAT 900  
QY 901 TCAATACAGACTTTTCTATGAGCTGAGATGCTTCAAGGTGATCTGCTTAACAGT 960  
DB 901 TCAATACAGACTTTTCTATGAGCTGAGATGCTTCAAGGTGATCTGCTTAACAGT 960  
QY 961 GATTCGAGGTAACAAGCTTACCTCAAGAGCTGAGCTTGAAGATTAAGAGAGAGCTGC 1020  
DB 961 GATTCGAGGTAACAAGCTTACCTCAAGAGCTGAGCTTGAAGATTAAGAGAGAGCTGC 1020  
QY 1021 GTCTTTTGAATAAAGGCAAGCAAGAAAGAAAGACTTACCTTACCCAGCATATA 1080  
DB 1021 GTCTTTTGAATAAAGGCAAGCAAGAAAGAAAGACTTACCTTACCCAGCATATA 1080  
QY 1081 CCTGCGGATGTTCTCTCAATTTTACCTGCTGCTTGAATTCGAGCAATTCCT 1140  
DB 1081 CCTGCGGATGTTCTCTCAATTTTACCTGCTGCTTGAATTCGAGCAATTCCT 1140  
QY 1141 AAAAGGCAATTTTGGAGAGCTTGGAGCTATACAGAGAGAGCTGAGAAAGGCAAG 1200  
DB 1141 AAAAGGCAATTTTGGAGAGCTTGGAGCTATACAGAGAGAGCTGAGAAAGGCAAG 1200  
QY 1201 CTACAGAGCTGAGAGTAACAAGAGGAGAGCGATTTATAGCGCTTTGTAAGAGAGC 1260  
DB 1201 CTACAGAGCTGAGAGTAACAAGAGGAGAGCGATTTATAGCGCTTTGTAAGAGAGC 1260  
QY 1261 TGTGCTGCTTGTGATCTCTCTGCTTCTGCTTCTGCTGCTGCTGCTGCTGCTG 1320  
DB 1261 TGTGCTGCTTGTGATCTCTCTGCTTCTGCTTCTGCTGCTGCTGCTGCTGCTG 1320  
QY 1321 CTGCTGGAACATTTCTTAACCTTCAACCAAGCAATTTCTGCTGAGAGCTTAAGTTA 1380  
DB 1321 CTGCTGGAACATTTCTTAACCTTCAACCAAGCAATTTCTGCTGAGAGCTTAAGTTA 1380  
QY 1381 TTTCACCGAAGAGCTTCTCAATTTCTCAATTTCTCAATTTCTCAATTTCTCAAT 1440  
DB 1381 TTTCACCGAAGAGCTTCTCAATTTCTCAATTTCTCAATTTCTCAATTTCTCAAT 1440  
QY 1441 ACAGAGGTTCTGCGAAGAGAGATGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500  
DB 1441 ACAGAGGTTCTGCGAAGAGAGATGATGAGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
QY 1501 CTTCAGCCAAACATATCATGATCCATGAAGACAGGAGGAGAGCCCTGAGCTTAAGATA 1560  
DB 1501 CTTCAGCCAAACATATCATGATCCATGAAGACAGGAGGAGAGCCCTGAGCTTAAGATA 1560  
QY 1561 TCCATCTCTCTGAGCAACAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
DB 1561 TCCATCTCTCTGAGCAACAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
QY 1621 ATTAATGTGGGTCCAGAAACCGGATAGCCCGTTTATGGGTTCTTACAACTAAGAG 1680  
DB 1621 ATTAATGTGGGTCCAGAAACCGGATAGCCCGTTTATGGGTTCTTACAACTAAGAG 1680  
QY 1681 AAACTCAGAAACAACCCAGATGGAATTTTGGAGCAATGTG--GTTTTTGGCTGC 1740  
DB 1681 AAACTCAGAAACAACCCAGATGGAATTTTGGAGCAATGTG--GTTTTTGGCTGC 1740  
QY 1741 AGGCAATGAAGATGAGGATTTCTATCAAGAAAGAGCTCAAGATTTCTTAAGCATGG 1800  
DB 1741 AGGCAATGAAGATGAGGATTTCTATCAAGAAAGAGCTCAAGATTTCTTAAGCATGG 1800  
QY 1801 ATCTTAATCATTAAGGTTCTCTCAAGAGATGCTCTGTTGGAGAGAGAGAGCC 1860  
DB 1801 ATCTTAATCATTAAGGTTCTCTCTCAAGAGATGCTCTCTGTTGGAGAGAGAGAGCC 1860  
QY 1861 CCAGCAAAAGTATGACAAAGCAATTCAGAGCTTATGAGCAGAGAGGAGAGATCTTC 1920  
DB 1861 CCAGCAAAAGTATGACAAAGCAATTCAGAGCTTATGAGCAGAGAGGAGAGATCTTC 1920  
QY 1921 CTCAGAGAGAGGCAATTTATGTGTGTGAGATGCAAGAAATATGCGCAAGATGTA 1980

DB 1918 CTCCAGAGAGAGGCAATTTATGTGTGAGATGCAAGAAATATGCGCAAGATGTA 1977  
QY 1981 CATGATGCCCTTGTGCAATATTAAGCAAGAGGTTGAGTTGAAGAACTAGAGCAATG 2040  
DB 1978 CATGATGCCCTTGTGCAATATTAAGCAAGAGGTTGAGTTGAAGAACTAGAGCAATG 2037  
QY 2041 AAAACCTGCGCACTTTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 2097  
DB 2038 AAAACCTGCGCACTTTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 2094

RESULT 6  
US-09-371-347-47  
; Sequence 47, Application US/09371347  
; Publication No. US20030082676A1  
; GENERAL INFORMATION:  
; APPLICANT: Roy A. Gravel et al.  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE REDUCTASE;  
; TITLE OF INVENTION: CLONING, AND METHODS FOR EVALUATING RISK OF NEURAL TUBE  
; TITLE OF INVENTION: DEFECTS, CARDIOVASCULAR DISEASE, AND CANCER  
; FILE REFERENCE: 50004/003003  
; CURRENT APPLICATION NUMBER: US/09/371,347  
; PRIOR FILING DATE: 1999-08-10  
; PRIOR APPLICATION NUMBER: 60/071,622  
; PRIOR FILING DATE: 1998-01-16  
; PRIOR APPLICATION NUMBER: 09/232,028  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 47  
; LENGTH: 2093  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-371-347-47

Query Match 85.5%; Score 1793; DB 10; Length 2093;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2093; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 1 ATGAGAGAGTTCTGTTACTATATGCTACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60  
DB 1 ATGAGAGAGTTCTGTTACTATATGCTACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60  
QY 61 GAAATGTGTGAGCAAGCTGTGTACATGATTTTCTGCAAGATCTTCACTGTATATGTA 120  
DB 61 GAAATGTGTGAGCAAGCTGTGTACATGATTTTCTGCAAGATCTTCACTGTATATGTA 120  
QY 121 TCCGATTAAGTATGACCTTAAACCGAAACAGCTCCCTGTTGTTGTGTTCTACACAG 180  
DB 121 TCCGATTAAGTATGACCTTAAACCGAAACAGCTCCCTGTTGTTGTGTTCTACACAG 180  
QY 181 GGCACCGGAGAGCCACCCGACAGAGCCCGCAAGTTGTTAAGAAATACAGAACCAACA 240  
DB 181 GGCACCGGAGAGCCACCCGACAGAGCCCGCAAGTTGTTAAGAAATACAGAACCAACA 240  
QY 241 CTGCGGTGATTTCTTGTCTCACTGCGGTATGAGTTACTGGGTCTCGGTATTCAGAA 300  
DB 241 CTGCGGTGATTTCTTGTCTCACTGCGGTATGAGTTACTGGGTCTCGGTATTCAGAA 300  
QY 301 TACACCTACTTTTGGCAATGGGGGAAAGTAATGATTAACGACTCAAGAGCTGAGGC 360  
DB 301 TACACCTACTTTTGGCAATGGGGGAAAGTAATGATTAACGACTCAAGAGCTGAGGC 360  
QY 361 CGGCAATTTCTATGACACTGAGATGATGATGATGATGATGATGATGATGATGATGATG 420  
DB 361 CGGCAATTTCTATGACACTGAGATGATGATGATGATGATGATGATGATGATGATGATG 420  
QY 421 CCGTGATGTTGCTGACTGTGCGCAAGCTTCAAGAAAGCAATTTTATGCTAGAGAGAGCA 480  
DB 421 CCGTGATGTTGCTGACTGTGCGCAAGCTTCAAGAAAGCAATTTTATGCTAGAGAGAGCA 480  
QY 481 GAGAGATTAAGTGGGCACTCCCGGTGATGACCTGATCCTTGAAGAGCAAGCTTGG 540

Db 481 GAGAGGTAAGTGGCGACTCCGGTGCATCACTCATCTTGTAGAGACAGACTTGTG 540  
Qy 541 AAGTCAGAGCTGTACATTCATTTGAATCTCAAGTCAGAGCTTGTAGATTCGATTCAGGA 600  
Db 541 AAGTCAGAGCTGTACATTCATTTGAATCTCAAGTCAGAGCTTGTAGATTCGATTCAGGA 600  
Qy 601 AGAAGAGATTCGTAGGTTTGAAGCAAAATGCACTGAACCAACCAATCCATGTTTGA 660  
Db 601 AGAAGAGATTCGTAGGTTTGAAGCAAAATGCACTGAACCAACCAATCCATGTTTGA 660  
Qy 661 ATTGAAGATTCGTAGGTTTGAAGCAAAATGCACTGAACCAACCAATCCATGTTTGA 720  
Db 661 ATTGAAGATTCGTAGGTTTGAAGCAAAATGCACTGAACCAACCAATCCATGTTTGA 720  
Qy 721 AATATTCCTGTAGTACCCCGAGATATTTACAGGTACATCTGCAGAGTCTCTTGGCCAG 780  
Db 721 AATATTCCTGTAGTACCCCGAGATATTTACAGGTACATCTGCAGAGTCTCTTGGCCAG 780  
Qy 781 GAGGAAGCCCAAGTATCTGTGACTTCAGCAGATCCAGTTTTCAGATTCGCAATTTCAAG 840  
Db 781 GAGGAAGCCCAAGTATCTGTGACTTCAGCAGATCCAGTTTTCAGATTCGCAATTTCAAG 840  
Qy 841 GCAGTTCAACTTACTAGAAATGATGCGATTAACCACTGCGGTGAGAAATGGACAT 900  
Db 841 GCAGTTCAACTTACTAGAAATGATGCGATTAACCACTGCGGTGAGAAATGGACAT 900  
Qy 901 TCAAAATCAGACTTTTCTTCACTAGCCTGAGAGTCCCTTCAAGCGTATCTCCCTTAACAT 960  
Db 901 TCAAAATCAGACTTTTCTTCACTAGCCTGAGAGTCCCTTCAAGCGTATCTCCCTTAACAT 960  
Qy 961 GATTCTGAGGTACAAAGCCTTACCAAGACTGCAAGTTTGAAGTAAGAGACACTGC 1020  
Db 961 GATTCTGAGGTACAAAGCCTTACCAAGACTGCAAGTTTGAAGTAAGAGACACTGC 1020  
Qy 1021 GTCCTTTGAAATTAAGGAGACACAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1080  
Db 1021 GTCCTTTGAAATTAAGGAGACACAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1080  
Qy 1081 CCGCGGAGATGTTCTCTCAGATTCATTTTACCTGTGTCTTGAATCCAGCAATTCCT 1140  
Db 1081 CCGCGGAGATGTTCTCTCAGATTCATTTTACCTGTGTCTTGAATCCAGCAATTCCT 1140  
Qy 1141 AAAAAGCATTTTTCGAGCCCTTGTGACTATACAGTACAGTGTCTGAAAAAGCCGAGG 1200  
Db 1141 AAAAAGCATTTTTCGAGCCCTTGTGACTATACAGTACAGTGTCTGAAAAAGCCGAGG 1200  
Qy 1201 CTAAGAGAGCTGTGAGTAAACAAAGGGGAGCGGATTAATAGCGCTTTTGAACAGATGCC 1260  
Db 1201 CTAAGAGAGCTGTGAGTAAACAAAGGGGAGCGGATTAATAGCGCTTTTGAACAGATGCC 1260  
Qy 1261 TGTGCTGTGTGTGATCTCTCTGCTGCTTCCCTTCTTGCCAGCAGCACTCAGTCTC 1320  
Db 1261 TGTGCTGTGTGTGATCTCTCTGCTGCTTCCCTTCTTGCCAGCAGCACTCAGTCTC 1320  
Qy 1321 CTGTGCAACATCTTCTTAAACTTCAACCCAGACCATATTCGTGTGCAAGCTCAAGTTTA 1380  
Db 1321 CTGTGCAACATCTTCTTAAACTTCAACCCAGACCATATTCGTGTGCAAGCTCAAGTTTA 1380  
Qy 1381 TTTCAACAGAGAAAGCTCAATTTGTCTCAACATTTGGAATTTCTGTCTACCTGCACA 1440  
Db 1381 TTTCAACAGAGAAAGCTCAATTTGTCTCAACATTTGGAATTTCTGTCTACCTGCACA 1440  
Qy 1441 ACAGAGTTCTGTGAGAGAGATATGACAGCTGTGCTGTGCTTGTGCTTCAAGT 1500  
Db 1441 ACAGAGTTCTGTGAGAGAGATATGACAGCTGTGCTGTGCTTGTGCTTCAAGT 1500  
Qy 1501 CTTGAGCCAAACATACATGATCCCATGAGAAGACGCGGAAAGCCCTGAGCTCTTAAGATA 1560  
Db 1501 CTTGAGCCAAACATACATGATCCCATGAGAAGACGCGGAAAGCCCTGAGCTCTTAAGATA 1560  
Qy 1561 TCCATCTCTCTGGAACAACAATTTCTTCACTTACCAATGACCCCTCAATCCCATC 1620  
Db 1561 TCCATCTCTCTGGAACAACAATTTCTTCACTTACCAATGACCCCTCAATCCCATC 1620

Db 1561 TCCATCTCTCTGGAACAACAATTTCTTCACTTACCAATGACCCCTCAATCCCATC 1620  
Qy 1621 ATATGAGTGGTTCAGGAACCGGCATAGCCCTTTATTTGGTTCCTACAACTAGAGAG 1680  
Db 1621 ATATGAGTGGTTCAGGAACCGGCATAGCCCTTTATTTGGTTCCTACAACTAGAGAG 1680  
Qy 1681 AAATCCAGAACACACCCAGATGGAATTTTGAAGCAATGTGTGTTTGTGCTGC 1740  
Db 1681 AAATCCAGAACACACCCAGATGGAATTTTGAAGCAATGTGTGTTTGTGCTGC 1740  
Qy 1741 AGGCATTAAGATGGAATTAATCTATTCAGAAAGACTCAGACATTTCTTAAGCATGG 1800  
Db 1741 AGGCATTAAGATGGAATTAATCTATTCAGAAAGACTCAGACATTTCTTAAGCATGG 1800  
Qy 1801 ATCTTAATCATTAAGGTTTCTTCTCAGAGATGCTCTGTGAGAGAGAGAGAGCC 1860  
Db 1801 ATCTTAATCATTAAGGTTTCTTCTCAGAGATGCTCTGTGAGAGAGAGAGAGCC 1860  
Qy 1861 CCAGCAAGTATGTACAAAGCAACATCCAGCTTCATGCGCAGAGTGGCGAATCTTC 1920  
Db 1861 CCAGCAAGTATGTACAAAGCAACATCCAGCTTCATGCGCAGAGTGGCGAATCTTC 1920  
Qy 1921 CTCAGAGAGAGCCCATATTTATGTGTGAGATGCAAAAGAAATAGCCAGAGATGTA 1980  
Db 1921 CTCAGAGAGAGCCCATATTTATGTGTGAGATGCAAAAGAAATAGCCAGAGATGTA 1980  
Qy 1981 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGTGTGAGTTGAAAACTAGAAAGCATG 2040  
Db 1981 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGTGTGAGTTGAAAACTAGAAAGCATG 2040  
Qy 2041 AAAACCTGGCCACTTTAAAGAAAGAAACGCTACCTTCAGATATTTGTCTATA 2097  
Db 2041 AAAACCTGGCCACTTTAAAGAAAGAAACGCTACCTTCAGATATTTGTCTATA 2097  
Db 2037 AAAACCTGGCCACTTTAAAGAAAGAAACGCTACCTTCAGATATTTGTCTATA 2093

RESULT 7  
US-10-741-600-692  
; Sequence 692, Application US/10741600  
; Publication No. US20050026169A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; FILE REFERENCE: W0004191  
; CURRENT APPLICATION NUMBER: US/10/741,600  
; CURRENT FILING DATE: 2003-12-22  
; NUMBER OF SEQ. ID NOS: 73997  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ. ID NO 692  
; LENGTH: 3256  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-741-600-692

Query Match 50.6%; Score 1062; DB 21; Length 3256;  
Best Local Similarity 99.1%; Pred. No. 0;  
Matches 2012; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 67 TGTGAGCAAGCTGTGTATACATGATTTTTCGAGATCTTCACTGATATTTAGTAATCCGAT 126  
Db 160 TGTGAGCAAGCTGTGTATACATGATTTTTCGAGATCTTCACTGATATTTAGTAATCCGAT 219  
Qy 127 AAGTATGACCTTAACCAAGCAAGCTCTCTGTGTGTGTGTTCTACAGGCGACC 186  
Db 220 AAGTATGACCTTAACCAAGCAAGCTCTCTGTGTGTGTGTTCTACAGGCGACC 279  
Qy 187 GGAAGCCACCCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACATGCGG 246  
Db 280 GGAAGCCACCCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACATGCGG 339  
Qy 247 GTTGATTTCTTTGCTACCTGCGGTATGGTTATCTGGGTCTCGGTATTCAGAAATCACC 306  
Db 340 GTTGATTTCTTTGCTACCTGCGGTATGGTTATCTGGGTCTCGGTATTCAGAAATCACC 399

Qy	307	TACTTTTGCAATGGGGGGGAATAATTAATGATAAAGACTTCAAGGCTTTGAAGCCCGGCAT	366
Db	400	TACTTTTGGCAATGGGGGGGAGATATTAATGATAAAGACTTCAAGGCTTTGAAGCCCGGCAT	459
Qy	367	TTCTATGACATCGAGCATGACAGATGACTGTGTAGGTTTGAACCTGTGTGTAGCCGTGG	426
Db	460	TTCTATGACACTCGAGCATATRCAGATGACTGTGTAGGTTTGAACCTGTGTGTAGCCGTGG	519
Qy	427	ATTGCTGAGACTCTGGCCAGACCCCTCAGAAAGCAATTTTAGTTCAGACAGAGCAAGAGAG	486
Db	520	ATTGCTGGAGCTCTGGCCAGACCCCTCAGAAAGCAATTTTAGTTCAGACAGAGCAAGAGAG	579
Qy	487	ATAAGTGCGCAGCTCCCGGTGGCATCACCTGCATCTTGAGGACAGACCTTGTGAATGCA	546
Db	580	ATAAGTGCGCAGCTCCCGGTGGCATCTCCTGCATCTTGAGGACAGACCTTGTGAATGCA	639
Qy	547	GAGGTGCTACACATGGAATCTCAAGTGGAGGCTTCGAGATTCGATGATTCAGGAAGAAG	606
Db	640	GAGGTGCTACACATGGAATCTCAAGTGGAGGCTTCGAGATTCGATGATTCAGGAAGAAG	699
Qy	607	GATTCGAGGTTTGAAGCAAAATGACGTGAACAGACAACAAATCAAATGTGTATTTGAA	666
Db	700	GATTCGAGGTTTGAAGCAAAATGACGTGAACAGACAACAAATCAAATGTGTATTTGAA	759
Qy	667	GACTTTGAGTCTGACATTAACCGGTCGGTACCCCACTGCACAAAGCCTCTGTAATATT	726
Db	760	GACTTTGAGTCTGACATTAACCGGTCGGTACCCCACTGCACAAAGCCTCTGTAATATT	819
Qy	727	CTGTGTTTACCCCGAGAAATATTACAGGTACATCTGCAGAGAGTCTTTGGCAGAGAGAA	786
Db	820	CTGTGTTTACCCCGAGAAATATTACAGGTACATCTGCAGAGAGTCTTTGGCAGAGAGAA	879
Qy	787	AGCCAAGATCTGTGACTTCAGACAGATCAGTTTTTCAAGGCCAATTTCAAGGACAGTT	846
Db	880	AGCCAAGATCTGTGACTTCAGACAGATCAGTTTTTCAAGGCCAATTTCAAGGACAGTT	939
Qy	847	CAACTTACTAGAATGATGATCATTAAAAACACTGTCTGTAGAATTGACATTTCAAT	906
Db	940	CAACTTACTAGAATGATGATCATTAAAAACACTGTCTGTAGAATTGACATTTCAAT	999
Qy	907	ACAGACTTTTCCCTATCAGCCTGGAGATGCTTCAAGCTGATCTGCCTACACATGATTTCT	966
Db	1000	ACAGACTTTTCCCTATCAGCCTGGAGATGCTTCAAGCTGATCTGCCTACACATGATTTCT	1059
Qy	967	GAGGTACAAACCCATCCCAAGACTGAGCTTGAAGTAAAGTAAAGAGCACTGGTCTT	1026
Db	1060	GAGGTACAAACCCATCCCAAGACTGAGCTTGAAGTAAAGTAAAGAGCACTGGTCTT	1119
Qy	1027	TTGAAAAATTAAGGAGACACAACAAGAAAGAGACTTACCCAGCATATACCTGGC	1086
Db	1120	TTGAAAAATTAAGGAGACACAACAAGAAAGAGACTTACCCAGCATATACCTGGC	1179
Qy	1087	GGATGTCTCTCAGTTCATTTTAACTGTGTCTTGAATCCGAGCAATTCCTAAAAAG	1146
Db	1180	GGATGTCTCTCAGTTCATTTTAACTGTGTCTTGAATCCGAGCAATTCCTAAAAAG	1239
Qy	1147	GCATTTTGGAGGCCCTTGTGACTATACAGTGAAGTCTGTAAAGCGCAGCTACAG	1206
Db	1240	GCATTTTTCAGGCCCTTGTGACTATACAGTGAAGTCTGTAAAGCGCAGCTACAG	1299
Qy	1207	GAGCTGTGACATTAACAAGGGGAGCGCAATTAAGCGCTTTGTACAGAGATGCTGTGGC	1266
Db	1300	GAGCTGTGACATTAACAAGGGGAGCGCAATTAAGCTGCTTTGTACAGAGATGCTGTGGC	1359
Qy	1267	TGCTTGTGTGATCTTCTCTCGCTTTCCTTCTTGCCAGCCACCACTCAGTCTCTGTCTC	1336
Db	1360	TGCTTGTGTGATCTTCTCTCGCTTTCCTTCTTGCCAGCCACCACTCAGTCTCTGTCTC	1419
Qy	1327	GAACATCTTCTTAAACTTCAACCCAGACCATATTCGTGTGCAAGTCAAGTTTATTTAC	1386
Db	1420	GAACATCTTCTTAAACTTCAACCCAGACCATATTCGTGTGCAAGTCAAGTTTATTTAC	1479

QY	1387	CCAGAGAAAGCCATTTTGGCTTCAACATTTGGAAATTTGCTACTGCACAACAG	1446
Db	1480	CCAGAGAAAGCTCCATTTTGTCTTCAACATTTGGAAATTTGCTACTGCACAACAG	1539
QY	1447	GTTCTGCGGAGAGAGATGTACACAGCTGGCTGGCTGTGTGGTTGCTCAGTTCCTCAG	1506
Db	1540	GTTCTGCGGAGAGAGATGTACACAGCTGGCTGGCTGTGTGGTTGCTCAGTTCCTCAG	1599
QY	1507	CCAAACATACATGTCATCCCATGAAGACAGCGGAAAGCCCTGGCTCTTAAGATATTCATC	1566
Db	1600	CCAAACATACATGTCATCCCATGAAGACAGCGGAAAGCCCTGGCTCTTAAGATATTCATC	1659
QY	1567	TCTCTCTGGAACAAATTTCTTTCCACTTACACATATACCCCTCAATCCCATATATATG	1626
Db	1660	TCTCTCTGGAACAAATTTCTTTCCACTTACACATATACCCCTCAATCCCATATATATG	1719
QY	1627	GTGGGTCCAGGACCGGATATGACCCCGTTATTTGGTTCTCTCAACACATAGAGAAATCTC	1686
Db	1720	GTGGGTCCAGGACCGGATATGACCCCGTTATTTGGTTCTCTCAACACATAGAGAAATCTC	1779
QY	1687	CAGAACAACAACCCAGATGTGAATTTTGGACATGTGGTTGTTTTTGGCTGACGCAT	1746
Db	1780	CAGAACAACAACCCAGATGTGAATTTTGGACATGTGGTTGTTTTTGGCTGACGCAT	1839
QY	1747	AAGATAGGGATTAATCTATTCAAGAAAGAGCTCAGACATTTCTTAAGCATGGATCTTA	1806
Db	1840	AAGATAGGGATTAATCTATTCAAGAAAGAGCTCAGATTTCTTAAGCATGGATCTTA	1899
QY	1807	ACTCATCTAAAGGTTTCTCTCTCAAGAGATGCTCTGTTGGGAGAGAGAACCCCAACA	1866
Db	1900	ACTCATCTAAAGGTTTCTCTCTCAAGAGATGCTCTGTTGGGAGAGAGAACCCCAACA	1959
QY	1867	AAGTATGTACAAGACAATCCACGCTTCATGGCAGAGGGGAGAGATCTCTCCAG	1926
Db	1960	AAGTATGTACAAGACAATCCACGCTTCATGGCAGAGGGGAGAGATCTCTCTCCAG	2019
QY	1927	GAGAACGGCCATATTTATGTGTGTGGAGATGCCAAGATATAGGCCTAAGATGTACATGAT	1986
Db	2020	GAGAACGGCCATATTTATGTGTGTGGAGATGCCAAGATATATAGGCCTAAGATGTACATGAT	2079
QY	1987	GCCCTTGTGCAATATATAGCAAGAGGTTGGATGAAAACTAGAAGCAATGAAAAAC	2046
Db	2080	GCCCTTGTGCAATATATAGCAAGAGGTTGGATGAAAACTAGAAGCAATGAAAAAC	2139
QY	2047	CTGGCCACTTTAAAGAGAAAAACGCTACCTTAAGATATTTGGTCATTA	2097
Db	2140	CTGGCCACTTTAAAGAGAAAAACGCTACCTTAAGATATTTGGTCATTA	2190

```

RESULT 8
US-10-741-600-693
; Sequence 693, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 693
; LENGTH: 3274
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-693

```

Query Match	50.6%	Score 1062;	DB 21;	Length 3274;
Best Local Similarity	99.1%;	Pred. No. 0;		
Matches 2012;	Conservative 0;	Mismatches 19;	Indels 0;	Gaps 0;

67 TGTGAGCAAGCTGTGGTACATGATGATTTTTCGACATCTTCACTGATATTAGTGAATCCGAT 126



FILE REFERENCE: AEOMICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 34288  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 6369  
LENGTH: 591  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC008727.5  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.45  
OTHER INFORMATION: NT HIT: AF121205.1, EVALUE 0.00e+00  
OTHER INFORMATION: SWISSPROT HIT: Q9UPW0, EVALUE 3.30e+00  
OTHER INFORMATION: EST\_HUMAN HIT: A0132586.1, EVALUE 0.00e+00  
US-10-029-386-6369

Query Match 15.7%; Score 330; DB 16; Length 591;  
Best Local Similarity 99.7%; Pred. No. 1.2e-169;  
Matches 380; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 401 GTTTAGAACTTGTGGTTGAGCCCTGATTCCTGCTGCGCCAGCCCTCAGAAAGCATT 460  
DB 38 GTTTAGAACTTGTGGTTGAGCCCTGATTCCTGCTGCGCCAGCCCTCAGAAAGCATT 97  
QY 461 TTAGGTCAGAGAGAGAGAGAGAGATTAAGTGGCGCACTCCCGGCGGATCAGCTGCAT 520  
DB 98 TTAGGTCAGAGAGAGAGAGAGAGATTAAGTGGCGCACTCCCGGCGGATCAGCTGCAT 157  
QY 521 CCTTGAAGACAGACCTTGTGAAGTCAAGCTGCTACATTAATCTCAAGTCAAGCTTC 580  
DB 158 CCTGAGAGACAGACCTTGTGAAGTCAAGCTGCTACATTAATCTCAAGTCAAGCTTC 217  
QY 581 TGAGATTCATGATTTTCAAGAGAAAGATTTCTGAGTTTGAAGCAAAATGCAGTGAAC 640  
DB 218 TGAGATTCATGATTTTCAAGAGAAAGATTTCTGAGTTTGAAGCAAAATGCAGTGAAC 277  
QY 641 GCAACCAATCCAAATGTTGATTAAGAGATTCGATCCCTCAGTACCCGTTGGGTACCC 700  
DB 278 GCAACCAATCCAAATGTTGATTAAGAGATTCGATCCCTCAGTACCCGTTGGGTACCC 337  
QY 701 CACTCTCACAAGCCTCTGTAATATTCCTGTTTACCCCAAGATATTTTACAGTACATC 760  
DB 338 CACTCTCACAAGCCTCTGTAATATTCCTGTTTACCCCAAGATATTTTACAGTACATC 397  
QY 761 TGCAGAGTCTCTTGGCCAGG 781  
DB 398 TGCAGAGTCTCTTGGCCAGG 418

RESULT 10  
US-10-029-386-20100  
Sequence 20100, Application US/10029386  
Publication No. US20030194704A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
FILE REFERENCE: AEOMICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 34288  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 20100  
LENGTH: 379  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC008727.5  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.45  
OTHER INFORMATION: NT HIT: g114729757, EVALUE 0.00e+00

OTHER INFORMATION: SWISSPROT HIT: Q9UPW0, EVALUE 1.80e+00  
OTHER INFORMATION: EST\_HUMAN HIT: A0132586.1, EVALUE 0.00e+00  
US-10-029-386-20100

Query Match 15.6%; Score 328; DB 16; Length 379;  
Best Local Similarity 99.7%; Pred. No. 1.5e-168;  
Matches 378; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 402 TTTAGAACTTGTGGTTGAGCCCTGATTCCTGCTGCGCCAGCCCTCAGAAAGCATT 461  
DB 1 TTTAGAACTTGTGGTTGAGCCCTGATTCCTGCTGCGCCAGCCCTCAGAAAGCATT 60  
QY 462 TAGTTCAGACAGAGACAGAGAGATTAAGTGGCGCACTCCCGTGGCATCCTGCATC 521  
DB 61 TAGTTCAGACAGAGACAGAGAGATTAAGTGGCGCACTCCCGTGGCATCCTGCATC 120  
QY 522 CTTGAAGACAGACCTTGTGAAGTCAAGCTGCTACATTAATCTCAAGTCAAGCTTCT 581  
DB 121 CTGAGAGACAGACCTTGTGAAGTCAAGCTGCTACATTAATCTCAAGTCAAGCTTCT 180  
QY 582 GAGATTCATGATTTTCAAGAGAAAGATTTCTAGGTTTGAAGCAAAATGCAGTGAACG 641  
DB 181 GAGATTCATGATTTTCAAGAGAAAGATTTCTAGGTTTGAAGCAAAATGCAGTGAACG 240  
QY 642 CAACCAATCCAAATGTTGATTAAGAGATTCGATCCCTCAGTACCCGTTGGGTACCC 701  
DB 241 CAACCAATCCAAATGTTGATTAAGAGATTCGATCCCTCAGTACCCGTTGGGTACCC 300  
QY 702 ACTTTCAGAGCCTCTGTAATATTCCTGTTTACCCCAAGATATTTTACAGTACATC 761  
DB 301 ACTTTCAGAGCCTCTGTAATATTCCTGTTTACCCCAAGATATTTTACAGTACATC 360  
QY 762 GCAGAGTCTCTTGGCCAG 780  
DB 361 GCAGAGTCTCTTGGCCAG 379

RESULT 11  
US-10-029-386-1735  
Sequence 1735, Application US/10029386  
Publication No. US20030194704A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
FILE REFERENCE: AEOMICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 34288  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 1735  
LENGTH: 591  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC021609.3  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.4  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.8  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.2  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.2  
OTHER INFORMATION: NT HIT: AF121205.1, EVALUE 0.00e+00  
OTHER INFORMATION: EST\_HUMAN HIT: A0132586.1, EVALUE 0.00e+00  
OTHER INFORMATION: SWISSPROT HIT: Q9UPW0, EVALUE 3.30e+00  
US-10-029-386-1735

Query Match 13.3%; Score 279; DB 16; Length 591;  
Best Local Similarity 99.5%; Pred. No. 1.2e-141;  
Matches 379; Conservative 0; Mismatches 2; Indels 0; Gaps 0;



QY 401 GTTTAGAACTTGTGGTGGACCGGTGATGCTGAGACTGTGGCCAGACCTTCAGAAAGCATT 460  
DB 38 GTTTAGAACTTGTGGTGGACCGGTGATGCTGAGACTGTGGCCAGACCTTCAGAAAGCATT 97  
QY 461 TTGAGTCAAGCAGAGAGCAAGAGAGATTAAGTGGCCGACCTCCGGTGGCATCCTGCAAT 520  
DB 98 TTGAGTCAAGCAGAGAGCAAGAGAGATTAAGTGGCCGACCTCCGGTGGCATCCTGCAAT 157  
QY 521 CCTTGAAGCAGACCTTGTGAAGCTCAGAGCTGTACATTTGATCTCAAGTGCAGCTTC 580  
DB 158 CCTGAGAGAGAGACCTGTGTAAGCTCAGAGCTGTACATTTGATCTCAAGTGCAGCTTC 217  
QY 581 TGAGATTGATGATTCAGAGAAAGAGATTCTGAGGTTTGAAGCAAAATGCACTGAACA 640  
DB 218 TGAGATTGATGATTCAGAGAAAGAGATTCTGAGGTTTGAAGCAAAATGCACTGAACA 277  
QY 641 GCAACCAATCCAAATGTTGTAATTGAAGACTTGAAGCTTCACTTACCCGTTGGTACCCC 700  
DB 278 GCAACCAATCCAAATGTTGTAATTGAAGACTTGAAGCTTCACTTACCCGTTGGTACCCC 337  
QY 701 CACTTCACAAGCCTCTGTAATTTCTGTTTACCCCAAGATTTTACAGGTACATC 760  
DB 338 CACTTCACAAGCCTCTGTAATTTCTGTTTACCCCAAGATTTTACAGGTACATC 397  
QY 761 TGCAGAGTCTCTTGGCCAGG 781  
DB 398 TGCAGAGTCTCTTGGCCAGG 418

RESULT 12  
US-10-029-386-15435  
; Sequence 15435, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:

; APPLICANT: Penn, Sharon G.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
; FILE REFERENCE: AEMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 3428  
; SOFTWARE: Annumax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 15435  
; LENGTH: 379  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC021609.3  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.4  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.8  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.2  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.2  
; OTHER INFORMATION: SWISSPROT HIT: Q9UPW0, EVALU0 1.80e+00  
; OTHER INFORMATION: EST HUMAN HIT: AU132586.1, EVALU0 0.00e+00  
; OTHER INFORMATION: NT HIT: g114729757, EVALU0 0.00e+00  
US-10-029-386-15435

Query Match 13.2%; Score 277; DB 16; Length 379;  
Best Local Similarity 99.5%; Pred. No. 1.5e-140;  
Matches 377; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 402 TTTTGAACCTTGTGGTGGACCGGTGATGCTGAGACTGTGGCCAGACCTTCAGAAAGCATT 461  
DB 1 TTTTGAACCTTGTGGTGGACCGGTGATGCTGAGACTGTGGCCAGACCTTCAGAAAGCATT 60  
QY 462 TAGGTCAAGCAGAGAGCAAGAGAGATTAAGTGGCCGACCTCCGGTGGCATCCTGATC 521  
DB 61 TAGGTCAAGCAGAGAGCAAGAGAGATTAAGTGGCCGACCTCCGGTGGCATCCTGATC 120

QY 522 CTTGAGAGACACCTTGTGAAGTCAAGAGCTGTACATTTGAATCTCAAGTGCAGCTTCT 581  
DB 121 CTTGAGAGAGACACCTTGTGAAGTCAAGAGCTGTACATTTGAATCTCAAGTGCAGCTTCT 180  
QY 582 GAGATTGATGATTCAGAGAAAGAGATTCTGAGGTTTGAAGCAAAATGCACTGAACA 641  
DB 181 GAGATTGATGATTCAGAGAAAGAGATTCTGAGGTTTGAAGCAAAATGCACTGAACA 240  
QY 642 CAACCAATCCAAATGTTGTAATTGAAGACTTGAAGCTTCACTTACCCGTTGGTACCCC 701  
DB 241 CAACCAATCCAAATGTTGTAATTGAAGACTTGAAGCTTCACTTACCCGTTGGTACCCC 300  
QY 702 ACTTCACAAGCCTCTGTAATTTCTGTTTACCCCAAGATTTTACAGGTACATC 761  
DB 301 ACTTCACAAGCCTCTGTAATTTCTGTTTACCCCAAGATTTTACAGGTACATC 360  
QY 762 GCAGAGTCTCTTGGCCAG 780  
DB 361 GCAGAGTCTCTTGGCCAG 379

RESULT 13  
US-10-741-600-17757  
; Sequence 17757, Application US/10741600  
; Publication No. US20050026169A1  
; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; FILE REFERENCE: CL001499  
; CURRENT APPLICATION NUMBER: US/10/741,600  
; CURRENT FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 73997  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17757  
; LENGTH: 43985  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-741-600-17757

Query Match 12.7%; Score 266; DB 21; Length 43985;  
Best Local Similarity 99.5%; Pred. No. 2.1e-134;  
Matches 366; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 401 GTTTAGAACTTGTGGTGGACCGGTGATGCTGAGACTGTGGCCAGACCTTCAGAAAGCATT 460  
DB 14836 GTTTAGAACTTGTGGTGGACCGGTGATGCTGAGACTGTGGCCAGACCTTCAGAAAGCATT 14895  
QY 461 TTGAGTCAAGCAGAGAGCAAGAGAGATTAAGTGGCCGACCTCCGGTGGCATCCTGCAAT 520  
DB 14896 TTGAGTCAAGCAGAGAGCAAGAGAGATTAAGTGGCCGACCTCCGGTGGCATCCTGCAAT 14955  
QY 521 CCTTGAAGCAGACCTTGTGAAGCTCAGAGCTGTACATTTGATCTCAAGTGCAGCTTC 580  
DB 14956 CCTTGAAGCAGACCTTGTGAAGCTCAGAGCTGTACATTTGATCTCAAGTGCAGCTTC 15015  
QY 581 TGAGATTGATGATTCAGAGAAAGAGATTCTGAGGTTTGAAGCAAAATGCACTGAACA 640  
DB 15016 TGAGATTGATGATTCAGAGAAAGAGATTCTGAGGTTTGAAGCAAAATGCACTGAACA 15075  
QY 641 GCAACCAATCCAAATGTTGTAATTGAAGACTTGAAGCTTCACTTACCCGTTGGTACCCC 700  
DB 15076 GCAACCAATCCAAATGTTGTAATTGAAGACTTGAAGCTTCACTTACCCGTTGGTACCCC 15135  
QY 701 CACTTCACAAGCCTCTGTAATTTCTGTTTACCCCAAGATTTTACAGGTACATC 760  
DB 15136 CACTTCACAAGCCTCTGTAATTTCTGTTTACCCCAAGATTTTACAGGTACATC 15195  
QY 761 TGCAGAG 768  
DB 15196 TGCAGAG 15203



RESULT 14  
US-10-029-386-633/c  
; Sequence 633, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO  
; FILE REFERENCE: AEOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 633  
; LENGTH: 525  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC021609.3  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.48  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.58  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.52  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.57  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.79  
; OTHER INFORMATION: SWISSPROT HIT: P37039, EVALUE 1.00e-06  
; OTHER INFORMATION: EST HUMAN HIT: BF346446.1, EVALUE 1.00e-98  
; OTHER INFORMATION: NT HIT: AF121212.1, EVALUE 0.00e+00  
US-10-029-386-633

Query Match 9 0%; Score 188; DB 16; Length 525;

Best Local Similarity 100.0%; Pred. No. 1.1e-91; Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1765 TTCAGAAAGAGCTCAGACATTTCTTAAGCATGGGATCTTAATCATTAAGGTTTC 1824  
DB 234 TTCAGAAAGAGCTCAGACATTTCTTAAGCATGGGATCTTAATCATTAAGGTTTC 175  
QY 1825 TTCTCAAGAGATGCTCTGTTGGGAGAGAAAGCCCAAGCAAGATGTACAGCAAC 1884  
DB 174 TTCTCAAGAGATGCTCTGTTGGGAGAGAAAGCCCAAGCAAGATGTACAGCAAC 115  
QY 1885 ATCCAGCTTCATGCGCAGAGGTTGGGAATCTCTCCAGAGAAAGCCCATTTAT 1944  
DB 114 ATCCAGCTTCATGCGCAGAGGTTGGGAATCTCTCCAGAGAAAGCCCATTTAT 55  
QY 1945 GTGTGTGG 1952  
DB 54 GTGTGTGG 47

RESULT 15

US-10-029-386-14338/c

; Sequence 14338, Application US/10029386

; Publication No. US20030194704A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharon G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G

; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO

; FILE REFERENCE: AEOMICA-X-2

; CURRENT APPLICATION NUMBER: US/10/029,386

; CURRENT FILING DATE: 2001-12-20

; NUMBER OF SEQ ID NOS: 34288

; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1

; SEQ ID NO 14338

; LENGTH: 175

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO AC021609.3

; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.48  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.58  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.52  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.57  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.79  
; OTHER INFORMATION: SWISSPROT HIT: O61608, EVALUE 4.00e-04  
; OTHER INFORMATION: EST HUMAN HIT: AA085543.1, EVALUE 7.00e-94  
; OTHER INFORMATION: NT HIT: G13325067, EVALUE 5.00e-94  
US-10-029-386-14338

Query Match 8 3%; Score 175; DB 16; Length 175;

Best Local Similarity 100.0%; Pred. No. 1.5e-84; Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1770 AAAAGGCTCAGACATTTCTTAAGCATGGGATCTTAATCATTAAGGTTTCCTTC 1829  
DB 175 AAAAGGCTCAGACATTTCTTAAGCATGGGATCTTAATCATTAAGGTTTCCTTC 116  
QY 1830 AAGAGTGCTCTGTTGGGAGAGAAAGCCCAAGCAAGATGTACAGCAACATCCA 1889  
DB 115 AAGAGTGCTCTGTTGGGAGAGAAAGCCCAAGCAAGATGTACAGCAACATCCA 56  
QY 1890 GCTTCATGCGCAGAGGTTGGGAATCTCTCCAGAGAAAGCCCATTTAT 1944  
DB 55 GCTTCATGCGCAGAGGTTGGGAATCTCTCCAGAGAAAGCCCATTTAT 1

Search completed: August 27, 2005, 17:33:22  
Job time : 903.401 secs

**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 23:18:31 ; Search time 4546.04 Seconds  
(without alignments)  
17558.328 Million cell updates/sec

Title: US-09-371-347A-41

Perfect score: 2097  
Sequence: 1 atgaggaggttcgtact.....ttcagatattgtgcataa 2097

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : EST.\*

- 1: gb\_est1.\*
- 2: gb\_est2.\*
- 3: gb\_hic.\*
- 4: gb\_est3.\*
- 5: gb\_est4.\*
- 6: gb\_est5.\*
- 7: gb\_est6.\*
- 8: gb\_gss1.\*
- 9: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1592	75.9	3100	3	BC062577 Homo sapi
2	1007	48.0	3143	3	BC035977 Homo sapi
3	740	35.3	874	4	BM801462 AGENCOURT
4	719	34.3	908	5	EX348674 BX348674
5	623	29.7	646	7	CN260357 170004241
6	586	27.9	852	5	BO431497 AGENCOURT
7	568	27.1	826	4	BI772430 603055786
8	565	26.9	565	1	AU279788 AU279788
9	543	25.9	877	1	AU124440 AU124440
10	531	25.3	1061	5	BQ218755 AGENCOURT
11	507	24.2	834	5	BU941078 AGENCOURT
12	470	22.4	521	6	CH164340 K-EST0225
13	461	22.0	776	6	CB997527 AGENCOURT
14	455	21.7	822	1	AU132586 AU132586
15	448	21.4	591	2	AW965709 EST377782
16	446	21.3	818	6	CD559384 AGENCOURT
17	434	20.7	591	4	BI025283 RCS-MT025
18	431	20.6	974	5	BX375211 BX375211
19	411	19.6	685	4	BM049352 603626120
20	406	19.4	710	5	BUS70323 AGENCOURT
21	384	18.3	527	7	BI025277 RCS-MT025
22	374	17.8	579	4	CN260360 170006001
23	367	17.5	642	2	BF346446 602020302
24	361	17.2	692	7	CN260359 170004706

25	359	17.1	499	6	CD704108 EST20635
26	354	16.9	386	1	AA279726 z92d10.r
27	351	16.7	839	4	BG531787 602560355
28	340	16.2	526	2	AW952883 EST364953
29	337	16.1	818	7	CP995233 AGENCOURT
30	335	16.0	413	2	BF810368 RCS-CI041
31	335	16.0	413	2	BF810479 RCS-CI014
32	332	15.8	366	1	AA085543 znu4h11.r
33	311	14.8	478	4	BM754488 K-EST0031
34	308	14.7	620	7	CK002453 AGENCOURT
35	308	14.7	664	7	CR768694 DKE2459K
36	308	14.7	667	7	CR770923 DKE2463N
37	308	14.7	767	7	CR557482 DKE2469K
38	302	14.4	440	4	BG877205 QV3-HT046
39	293	14.0	481	7	CR549172 DKE2459J
40	292	13.9	528	2	BE301292 ba8pb07.x
41	290	13.8	395	4	BM838530 K-EST0114
42	272	13.0	301	1	AL704780 DKE2468M
43	264	12.6	366	6	CB298361 220019.re
44	257	12.3	366	2	BF808461 QV1-CI017
45	257	12.3	368	1	AA355001 EST63417

## ALIGNMENTS

RESULT 1	BC062577	3100 bp	mRNA	linear	HTC 25-NOV-2003
LOCUS	BC062577				
DEFINITION	Homo sapiens cDNA clone IMAGE:5189058, containing frame-shift errors.				
ACCESSION	BC062577.1	GI:38511756			
VERSION					
KEYWORDS	HTC.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Strausberg, R.D., Collins, F.S., Wagner, L.H., Derge, J.G., Klausner, R.D., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bat, N.K., Altschul, S.F., Moore, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Hopkins, R.F., Jordan, H., Farmer, A.A., Rubin, G.M., Hong, L., Ditschenko, L., Marusina, K., Bonaldo, M.F., Casavant, T.L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uebelin, T.B., Toshitani, S., Carninci, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McMan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hilyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, U., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywicki, M.I., Skalek, U., Small, D.B., Scherker, A., Schein, J.E., Jones, S.J., and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
TITLE	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
JOURNAL	22388257				
MEDLINE	12477932				
PUBMED	2 (bases 1 to 3100)				
REFERENCE	Strausberg, R.				
AUTHORS	Direct Submission				
TITLE	Submitted (24-NOV-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,				
JOURNAL	USA NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>				
REMARK	Contact: MGC help desk				
COMMENT	Email: <a href="mailto:cgabbs-r@mail.nih.gov">cgabbs-r@mail.nih.gov</a>				
	Tissue Procurement: Life Technologies, Inc.				

cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: nisc.ncm@nih.gov  
Ahter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
Blakesley, R.W., Boufield, G.G., Breen, K., Brinkley, C., Brooks, S.,  
Dietrich, N.L., Graniter, S., Guan, X., Gupta, J., Haghighi, P.,  
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,  
Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,  
McDowell, J., Pearson, R., Stancirip, S., Thomas, P.J., Touchman, J.W.,  
Tsungeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LIML at: <http://image.liml.gov>  
Series: IRAC Plate: 135 Row: e Column: 21  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA 91: 4505278  
This clone has the following problem: frame shifted.

## FEATURES

## source

1. 3100  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5189058"  
/tissue\_type="Colon, Kidney, Stomach, adult, whole pooled"  
/clone\_lib="NIH MGC\_116"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"

## ORIGIN

Query Match 75.9%; Score 1592; DB 3; Length 3100;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1812; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
283 GGTCTCGGTGATTCAGATACACCTTCTTTCAGATGAGGAGGAGATTAATGATTAACGA 342  
172 GGTCTCGGTGATTCAGATACACCTTCTTTCAGATGAGGAGGAGATTAATGATTAACGA 231  
343 CTTCAAGAGCTTGAAGCCCGGCACTTTCTATGACACTGACATGACATGACATGATGATGAT 402  
232 CTTCAAGAGCTTGAAGCCCGGCACTTTCTATGACACTGACATGACATGACATGATGATGAT 291  
403 TTGAACTTGTGTTGAGCGGTGATGCTGACCTGCGCAGCCCTCAGAAAGCATTTT 462  
292 TTGAACTTGTGTTGAGCGGTGATGCTGACCTGCGCAGCCCTCAGAAAGCATTTT 351  
463 AGGTCAAGCAGAGCAGAGAGAGATTAAGTGGCGCACTCCCGGTGACATCACTGCATCC 522  
352 AGGTCAAGCAGAGCAGAGAGAGATTAAGTGGCGCACTCCCGGTGACATCACTGCATCC 411  
523 TTGAGGACAGACTTGTGAAGTCAAGCTGCTACATTTGAATTTCAATGACATTTCTG 582  
412 TCGAGGACAGACTTGTGAAGTCAAGCTGCTACATTTGAATTTCAATGACATTTCTG 471  
583 AGATTGATGATTCAGAGAGAGAGATTCAGAGTTTGAAGCAAAATGACAGTGAAGC 642  
472 AGATTGATGATTCAGAGAGAGAGATTCAGAGTTTGAAGCAAAATGACAGTGAAGC 531  
643 AACCAATTCATGTTGTAATTTGAAGCTTGAAGTCTTCACTTACCCGTTGCGTACCCCA 702  
532 AACCAATTCATGTTGTAATTTGAAGCTTGAAGTCTTCACTTACCCGTTGCGTACCCCA 591  
703 CTCTCAAGAGCTTCTGAATTTCTGAGTTTACCCCAAGATTTTACAGATTCAGATTCAG 762  
592 CTCTCAAGAGCTTCTGAATTTCTGAGTTTACCCCAAGATTTTACAGATTCAGATTCAG 651  
763 CAGAGTCTCTTGGCAGAGAGAGAGCAAGTATCTGATGATTCAGAGATTCAGATTCAG 822  
652 CAGAGTCTCTTGGCAGAGAGAGAGCAAGTATCTGATGATTCAGAGATTCAGATTCAG 711

QY 823 CAAATGCCAATTTCAAGGAGAGTCACTTACTAGATGATGCCATTAACCACTCTG 882  
DB 712 CAAATGCCAATTTCAAGGAGAGTCACTTACTAGATGATGCCATTAACCACTCTG 771  
QY 883 CAGTGAATTTGAGCACTTCAAAATCAGACTTTCTATCAGCTGAGATGCTTCAG 942  
DB 772 CTGTGAGAAATGAGCACTTCAAAATCAGACTTTCTATCAGCTGAGATGCTTCAG 831  
QY 943 GTGATGCTGCTTAACAGTATTTCTGAGTACAAAGCTTCTCCAAAGCTGACGTTGAA 1002  
DB 832 GTGATGCTGCTTAACAGTATTTCTGAGTACAAAGCTTCTCCAAAGCTGACGTTGAA 891  
QY 1003 GATTAAGAGAGCACTGCGTCTTTGAAATTAAGGAGAGCAAGAGAGAGAGAGCT 1062  
DB 892 GATTAAGAGAGCACTGCGTCTTTGAAATTAAGGAGAGCAAGAGAGAGAGAGCT 950  
QY 1063 ACCTTACCCAGATATATCTGAGGAGATTTCTTCCATTCATTTTACCTGCTCTT 1122  
DB 951 ACCTTACCCAGATATATCTGAGGAGATTTCTTCCATTCATTTTACCTGCTCTT 1010  
QY 1123 GAAATCCGAGCAATTTCTTAAGAGAGATTTTTCAGAGCTTGTGACTATACAGTGA 1182  
DB 1011 GAAATCCGAGCAATTTCTTAAGAGAGATTTTTCAGAGCTTGTGACTATACAGTGA 1070  
QY 1183 AGTCTGAAG 1242  
DB 1071 AGTCTGAAG 1130  
QY 1243 CGCTTTGTCAG 1302  
DB 1131 CGCTTTGTCAG 1190  
QY 1303 CAGCAGCACTAGTCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1362  
DB 1191 CAGCAGCACTAGTCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1250  
QY 1363 TGTGCAAGCTCAAGTTTATTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1422  
DB 1251 TGTGCAAGCTCAAGTTTATTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1310  
QY 1423 TTTCTGCTACAG 1482  
DB 1311 TTTCTGCTACAG 1370  
QY 1483 TTGTTGTTGCTTCAAGTTTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1542  
DB 1371 TTGTTGTTGCTTCAAGTTTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1430  
QY 1543 GCTCTGCTTCAAGATATTCATCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1602  
DB 1431 GCTCTGCTTCAAGATATTCATCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1490  
QY 1603 GAGCCTGCTTCAAGATATTCATCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1662  
DB 1491 GAGCCTGCTTCAAGATATTCATCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1550  
QY 1663 TTCTCAACAATAG 1722  
DB 1551 TTCTCAACAATAG 1610  
QY 1723 TGGTTGTTTTTGGCTGAG 1782  
DB 1611 TGGTTGTTTTTGGCTGAG 1670  
QY 1783 CATTTGCTTAAG 1842  
DB 1671 CATTTGCTTAAG 1730  
QY 1843 GTTGGGAG 1902  
DB 1731 GTTGGGAG 1790

Qy	1903	CAGGTGGCGAAATCCCTCCACAGAGAACGGCCATATTTATTTGTGTGGAGATGCAAG	1962
Db	1791	CAGGTGGCGAAATCCCTCCACAGAGAACGGCCATATTTATTTGTGTGGAGATGCAAG	1850
Qy	1963	AATATGGCCAAAGATGTATCATATGATGACCTTGTGTGCAAAATATATACAAAGAGTTGAGTT	2022
Db	1851	AATATGGCCAAAGATGTATCATATGATGACCTTGTGTGCAAAATATATACAAAGAGTTGAGTT	1910
Qy	2023	GAATAACTAGAAAGCATATGAAAAACCTGGCCACTTTAAAAAGAGAAAAACGTAACCTTCAG	2082
Db	1911	GAATAACTAGAAAGCATATGAAAAACCTGGCCACTTTAAAAAGAGAAAAACGTAACCTTCAG	1970
Qy	2083	GATATTTGGTCATTA	2097
Db	1971	GATATTTGGTCATTA	1985
RESULT 2			
LOCUS	BC035977	3143 bp	mRNA
DEFINITION	Homo sapiens, clone IMAGE:4611253,	linear	HTC 20-SEP-2002
ACCESSION	BC035977		
VERSION	BC035977.1	GI:23243305	
KEYWORDS	HTC.		
SOURCE	Homo sapiens (human)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
REFERENCE	1 (bases 1 to 3143)		
AUTHORS	Strausberg,R.		
TITLE	Direct Submission		
JOURNAL	Submitted (31-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>		
COMMENT	Contact: MGC help desk Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a> Tissue Procurement: CLONTECH cDNA Library Preparation: CLONTECH Laboratories, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: <a href="http://www.shgc.stanford.edu">http://www.shgc.stanford.edu</a> Contact: (Dickson, Mark) <a href="mailto:mcdpaxil.stanford.edu">mcdpaxil.stanford.edu</a> Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.		
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Series: IRAL Plate: 41 Row: g Column: 2 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4505278 This clone has the following problem: frame shifted.		
SOURCE	Location/Qualifiers		
	1. 3143		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:4611253"		
	/tissue_type="Kidney"		
	/clone_id="NIH MGC_75"		
	/lab_host="DH10B"		
	/note="Vector: pDNR-LIB"		
ORIGIN			
Query Match	48.0%;	Score 1007;	DB 3; Length 3143;
Best Local Similarity	99.9%;	Pred. No. 0;	
Matches 1057;	Conservative 0;	Mismatches 1;	Indels 0; Gaps 0;
Qy	1	ATGAGAGAGTTTCTGTACTATATCTACACAGCAGGAGCAAGGCAAGCCATGCGAGAA	60
Db	52	ATGAGAGAGTTTCTGTACTATATCTACACAGCAGGAGCAAGGCAAGCCATGCGAGAA	111

Qy	61	GAATATATGTAAGCAAGCTGTGTGATCAATGGAATTTTCTGCAATCTGATCACTGTAATTAAGAA	120
Db	112	GAATATATGTAAGCAAGCTGTGTGATCAATGGAATTTTCTGCAATCTGATCACTGTAATTAAGAA	171
Qy	121	TCCGATATAGTATGACCTTAATAAACCGAAACAGCTCTCTTGTTGTTGTGTGTTCTTACACG	180
Db	172	TCCGATATAGTATGACCTTAATAAACCGAAACAGCTCTCTTGTTGTTGTGTGTTCTTACACG	231
Qy	181	GGACCCGAGAACCCACCCGACACAGCCCGCAAGTTTGTTAAGAAATACAGAACCAACA	240
Db	232	GGACCCGAGAACCCACCCGACACAGCCCGCAAGTTTGTTAAGAAATACAGAACCAACA	291
Qy	241	CTGCCGGTGAATTTCTTGTCTCACCTGCGGTATAGGGTTACTGCGGTCTCGGATTCAGAA	300
Db	292	CTGCCGGTGAATTTCTTGTCTCACCTGCGGTATAGGGTTACTGCGGTCTCGGATTCAGAA	351
Qy	301	TACACCTACTTTTGCATAGGGGGGGAAGATATATTGATTAACGACTTCAAGACTTGGAGCC	360
Db	352	TACACCTACTTTTGCATAGGGGGGGAAGATATATTGATTAACGACTTCAAGACTTGGAGCC	411
Qy	361	CGGCATTTCTATGACATGCAATGCAATGACTGTGTAGGTTTGAATTGTGTGTGAG	420
Db	412	CGGCATTTCTATGACATGCAATGCAATGACTGTGTAGGTTTGAATTGTGTGTGAG	471
Qy	421	CCGGTGAATGCTGGAATCTGCGCCAGCCCTCAGAAAGATTTTATGTAAGCAGAGGACAA	480
Db	472	CCGGTGAATGCTGGAATCTGCGCCAGCCCTCAGAAAGATTTTATGTAAGCAGAGGACAA	531
Qy	481	GAGAGATTAAGTGGCGGACTCCCGGTGGCATTCACCTGCACTCTTGAAGACAGACTTGTG	540
Db	532	GAGAGATTAAGTGGCGGACTCCCGGTGGCATTCACCTGCACTCTTGAAGACAGACTTGTG	591
Qy	541	AAGTCAAGCTGTACACATTTGAATCTCAAGTGAAGCTTCTGAGATTGATGATTCAGGA	600
Db	592	AAGTCAAGCTGTACACATTTGAATCTCAAGTGAAGCTTCTGAGATTGATGATTCAGGA	651
Qy	601	AGAAAGATTTCTGAGTTTGAAGCAAAATGSCATGAAACAGCAACCAATCCAATGTTGTA	660
Db	652	AGAAAGATTTCTGAGTTTGAAGCAAAATGSCATGAAACAGCAACCAATCCAATGTTGTA	711
Qy	661	ATTGAAGACTTTGAGTCTCTCACTTACCCTGTCGGTACCCCACTCTCAACAAGCTCTCTG	720
Db	712	ATTGAAGACTTTGAGTCTCTCACTTACCCTGTCGGTACCCCACTCTCAACAAGCTCTCTG	771
Qy	721	AATATTCCTGTTTACCCTCCAGAAATATTTACAGTACATCTGCAGAGTCTTGTGGCCAG	780
Db	772	AATATTCCTGTTTACCCTCCAGAAATATTTACAGTACATCTGCAGAGTCTTGTGGCCAG	831
Qy	781	GAGAAAGCCCAAGTATCTGTGACTTGCAGATCCAGATTCAGGTTTCAAGGCCAATTTCAAAG	840
Db	832	GAGAAAGCCCAAGTATCTGTGACTTGCAGATTCAGGTTTCAAGGCCAATTTCAAAG	891
Qy	841	GCACTTCACTTACTACGAATGATGACCATTAATAAACCACTCTGCTGTGATTAAGACATT	900
Db	892	GCACTTCACTTACTACGAATGATGACCATTAATAAACCACTCTGCTGTGATTAAGACATT	951
Qy	901	TCAAATACAGACTTTTCCATACGCTTGAGAGATGCTTCAAGCTGATCTGCTTACACAT	960
Db	952	TCAAATACAGACTTTTCCATACGCTTGAGAGATGCTTCAAGCTGATCTGCTTACACAT	1011
Qy	961	GATTCTGAGTACAAAGCTATCTCAAAAGCTGCAAGCTGTAAGATTAATAAGAGACACATGC	1020
Db	1012	GATTCTGAGTACAAAGCTATCTCAAAAGCTGCAAGCTGTAAGATTAATAAGAGACACATGC	1071
Qy	1021	GTCTCTTTGAAATTAAGCAGACACAAAGAAAGAG	1058
Db	1072	GTCTCTTTGAAATTAAGCAGACACAAAGAAAGAG	1109
RESULT 3			
BM801462			
LOCUS			
BM801462 874 bp mRNA linear EST 05-MAR-2002			

```

DEFINITION AGENCOURT_6459212 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5560477
5', mRNA sequence.
ACCESSION BM801462
VERSION BM801462.1 GI:19118285
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 874)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
AUTHORS Unpublished (1999)
TITLE Contact: Robert Strausberg, Ph.D.
JOURNAL Email: sgabs-remail.nih.gov
COMMENT Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM12286 row: 1 column: 14
High quality sequence stop: 710.
Location/Qualifiers
1..874
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5560477"
/libsize_type="duodenal adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_88"
/note="Organ: small intestine; Vector: pCMV-SPORTe;
Site 1: NotI; Site 2: SalI; Cloned unidirectionally;
oligo-dT primed. Average insert size 1.767 kb. Library
enriched for full-length clones and constructed by Life
Technologies. Note: this is a NIH_MGC Library."
ORIGIN
Query Match 35.3%; Score 740; DB 4; Length 874;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGAGAGTTTCTGTTACTATATGCTACACAGCAGGAGACAGCAAGGCATGCGAGAA 60
DB 50 ATGAGAGAGTTTCTGTTACTATATGCTACACAGCAGGAGACAGCAAGGCATGCGAGAA 109
QY 61 GAATATATGTGAGCAAGCTGTGTGATCATGAGATTCTGCAATCTTCACTGTATTAGTAA 120
DB 110 GAATATATGTGAGCAAGCTGTGTGATCATGAGATTCTGCAATCTTCACTGTATTAGTAA 169
QY 121 TCGGATATGATGACCTTAATAACGAAACAGCTCCTGTGTGTGTGTTGTTCTTACACAG 180
DB 170 TCGGATATGATGACCTTAATAACGAAACAGCTCCTGTGTGTGTGTTGTTCTTACACAG 229
QY 181 GGACCCGAGAGCCACCCGACACAGCCCGCAAGTTTGTTAAGAAATACAGAACCAACA 240
DB 230 GGACCCGAGAGCCACCCGACACAGCCCGCAAGTTTGTTAAGAAATACAGAACCAACA 289
QY 241 CTCGCGGTTGATTTCTTGTCTCACTCGGAGTATGAGGTTCTCGGTTCTCGGTTAGAA 300
DB 290 CTCGCGGTTGATTTCTTGTCTCACTCGGAGTATGAGGTTCTCGGTTCTCGGTTAGAA 349
QY 301 TACACCTACTTTGCAATGGGGGGAAGATATGTATAAGACTTCAAGAGCTTGGAGCC 360
DB 350 TACACCTACTTTGCAATGGGGGGAAGATATGTATAAGACTTCAAGAGCTTGGAGCC 409
QY 361 CGGCAATTTCTATGACATGACATGACATGACATGATGATGATTTAGAACTTGTGTTAG 420
DB 410 CGGCAATTTCTATGACATGACATGACATGACATGATGATGATTTAGAACTTGTGTTAG 469
QY 421 CGTGATTTGCTGAGACTGTGGCCAGCCCTCAGAAAGCATTTTAGGTCAAGCAGAGCAAA 480

```

```

DB 470 CCGTGATTTGCTGAGACTGTGGCCAGCCCTCAGAAAGCATTTTAGGTCAAGCAGAGCAAA 529
QY 481 GAGAGATATAGTGGCGCACTCCCGGTGGCATCACTGATCTCTTGAAGACAGACCTTGTG 540
DB 530 GAGAGATATAGTGGCGCACTCCCGGTGGCATCACTGATCTCTTGAAGACAGACCTTGTG 589
QY 541 AAGTCAGACCTGCTACACATTTGATATTCAGATGACCTTCTGAGATTTGATGAGA 600
DB 590 AAGTCAGACCTGCTACACATTTGATATTCAGATGACCTTCTGAGATTTGATGAGA 649
QY 601 AGAAGACTTTGAGGTTTGAAGCAAAATGACAGTGAACAGCAACCATTCATGTTGTA 660
DB 650 AGAAGACTTTGAGGTTTGAAGCAAAATGACAGTGAACAGCAACCATTCATGTTGTA 709
QY 661 ATTGAAGCTTTGAGTCTCATCTTACCTCCGTGGTACCCCACTCTCAGAACCTCTCG 720
DB 710 ATTGAAGCTTTGAGTCTCATCTTACCTCCGTGGTACCCCACTCTCAGAACCTCTCG 769
QY 721 AATATTCCTGTTTACCCCC 740
DB 770 AATATTCCTGTTTACCCCC 789

RESULT 4
BX348674 908 bp mRNA linear EST 08-APR-2004
LOCUS BX348674 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
DEFINITION CDNA clone CS00C010Y11 5-PRIME, mRNA sequence.
ACCESSION BX348674
VERSION BX348674.1 GI:30375301
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 908)
AUTHORS Li, W.-B., Gruber, C., Jessee, J. and Polyes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: segre@genoscope.cns.fr, web : www.genoscope.cns.fr
Note: segre@genoscope.cns.fr, web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
3392.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?b=CS0BAG062B02_CS00490_1&c=3392.f
FEATURES
source Location/Qualifiers
1..908
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0C010Y11"
/libsize_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA
was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 34.3%; Score 719; DB 5; Length 908;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 719; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 674 AGTCTCACTTACCGTGGTACCCCACTCAGCAAGCTCTCGAATATTCGGTT 733

```

```

Db      28 AGTCTCACTACCCGTTGGTACCCCACTCTCAACAGCTCTGTGAATATTCCTGT 87
Qy      734 TACCCCAATATTTTACAGTACATCTTCAGAGAGCTTTGGCCAGGAGAAAGCAAG 793
Db      88 TACCCCAATATTTTACAGTACATCTTCAGAGAGCTTTGGCCAGGAGAAAGCAAG 147
Qy      794 TATCTGTGACTTCAGAGATCCAGTTTTCAGTGCCATTTCAAGGAGTTCACTTA 853
Db      148 TATCTGTGACTTCAGAGATCCAGTTTTCAGTGCCATTTCAAGGAGTTCACTTA 207
Qy      854 CTACAGATGATCCATTAATAACCACTCTGCTGTAGATTTGACATTTCAATACAGCT 913
Db      208 CTACAGATGATCCATTAATAACCACTCTGCTGTAGATTTGACATTTCAATACAGCT 267
Qy      914 TTTCCATACGCTTGAGATGCTTTCAGGCTATCTGCTTAAAGTATTTCTGAGTAC 973
Db      268 TTTCCATACGCTTGAGATGCTTTCAGGCTATCTGCTTAAAGTATTTCTGAGTAC 327
Qy      974 AAAGCTTCTCCAAAGCTGAGCTTGAAGATTAAGAGAGCACTGCTCTTTGAAA 1033
Db      328 AAAGCTTCTCCAAAGCTGAGCTTGAAGATTAAGAGAGCACTGCTCTTTGAAA 387
Qy      1034 TAAAGGACAGACAAAGAAAGAAAGAGCTTACCTTACCCAGCATATACCTGGGAGT 1093
Db      388 TAAAGGACAGACAAAGAAAGAAAGAGCTTACCTTACCCAGCATATACCTGGGAGT 447
Qy      1094 CTCTCCAGTTCAATTTTACCTGTGTCTTGAATTCGAGCAATCTTAAAGGCAATTT 1153
Db      448 CTCTCCAGTTCAATTTTACCTGTGTCTTGAATTCGAGCAATCTTAAAGGCAATTT 507
Qy      1154 TCGAGAGCTTGTGAGCTATACAGAGAGAGCTTAAAGGAGGAGCTTACAGAGAGCT 1213
Db      508 TCGAGAGCTTGTGAGCTATACAGAGAGAGCTTAAAGGAGGAGCTTACAGAGAGCT 567
Qy      1214 GCAGTAAACAGAGGAGAGCGATATATAGCGCTTTGTACAGATGCTGTGCTGTGT 1273
Db      568 GCAGTAAACAGAGGAGAGCGATATATAGCGCTTTGTACAGATGCTGTGCTGTGT 627
Qy      1274 TGGATCTCTCTCTGCTTCCCTTCTTTCGACGACCACTCACTCTCTCTGCTGAACT 1333
Db      628 TGGATCTCTCTCTGCTTCCCTTCTTTCGACGACCACTCACTCTCTCTGCTGAACT 687
Qy      1334 TTCCTAACTTCAACCCAGACCATATTCGTGTGCAAGCTCAAGTTATTTCAACCCAGA 1392
Db      688 TTCCTAACTTCAACCCAGACCATATTCGTGTGCAAGCTCAAGTTATTTCAACCCAGA 746

```

```

RESULT 5
CN260357      646 bp      mRNA      linear      EST 16-MAY-2004
DEFINITION    17000424179730 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
ACCESSION     CN260357
VERSION       CN260357.1 GI:47276771
KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)

```

```

REFERENCE
AUTHORS       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
              1 (bases 1 to 646)
              Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murgue, J., Fisk, G.J.,
              Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
              Lebkoweki, J and Stanton, L.W.
              Transcriptome characterization elucidates signaling networks that
              control human ES cell growth and differentiation
              Nat. Biotechnol. 22 (6), 707-716 (2004)
              Contact: Brandenberger R
              Regenerative Medicine
              Geron Corporation
              230 Constitution Drive, Menlo Park, CA 94025, USA
              Tel: 650 473 8658
              Fax: 650 473 7760
              Email: rbrandenberger@geron.com
              Insert Length: 646 Std Error: 0.00.

```

## FEATURES

source

Location/Qualifiers

```

1..646
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, cell lines H1, H7, and
H9"
/clone_1ib="GRN ES"
/note="oligo dt primed, full-length enriched cDNA library
from undifferentiated hES cell lines H1 (p32), H7 (p29),
and H9 (p26) maintained in feeder-free conditions"

```

## ORIGIN

Query Match 29.7%; Score 623; DB 7; Length 646;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 623; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      987 AAGACTGACAGCTTGAAGATTAAGAGAGAGCACTGCTCTTTGAAAATTAAGGACAGAC 1046
Db      24 AAGACTGACAGCTTGAAGATTAAGAGAGAGCACTGCTCTTTGAAAATTAAGGACAGAC 83
Qy      1047 AAAGAAAGAAAGAGCTTACCTTACCCAGATATACCTGGGAGATGTTCTCTCAATCAT 1106
Db      84 AAAGAAAGAAAGAGCTTACCTTACCCAGATATACCTGGGAGATGTTCTCTCAATCAT 143
Qy      1107 TTTTACCTGCTGTCTTGAATTCGAGCAATTCCTAATAAGGCAATTTTTCGAGCCCTGT 1166
Db      144 TTTTACCTGCTGTCTTGAATTCGAGCAATTCCTAATAAGGCAATTTTTCGAGCCCTGT 203
Qy      1167 GGAATATACAGATGACAGTGTGAAAAGGAGGAGCTTACAGAGCTGTGAGTAAACAG 1226
Db      204 GGAATATACAGATGACAGTGTGAAAAGGAGGAGCTTACAGAGCTGTGAGTAAACAG 263
Qy      1227 GGAAGCCGATTTATAGCCGCTTTGTACAGATGCTGTGCTGTGTGTGTGTCTCT 1286
Db      264 GGAAGCCGATTTATAGCCGCTTTGTACAGATGCTGTGCTGTGTGTGTGTGTCTCTCT 323
Qy      1287 GCGTTCCCTTCTTTCGAGGACCACTGCTGCTGCTGGAACATCTTCTTAACTTCA 1346
Db      324 GCGTTCCCTTCTTTCGAGGACCACTGCTGCTGCTGGAACATCTTCTTAACTTCA 383
Qy      1347 ACCAGAACCATATTCGTGTGACAGCTCAAGTTATTTACCAAGAAAGCTCATTTGT 1406
Db      384 ACCAGAACCATATTCGTGTGACAGCTCAAGTTATTTACCAAGAAAGCTCATTTGT 443
Qy      1407 CTTCACATTTGTGAATTTCTGTCTTACTGCAACACAGAGTTCTGCGAAGGAGTATG 1466
Db      444 CTTCACATTTGTGAATTTCTGTCTTACTGCAACACAGAGTTCTGCGAAGGAGTATG 503
Qy      1467 TACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1526
Db      504 TACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 563
Qy      1527 TGAAGACAGCGGAGAAAGCCCTGCTCTTAAGATTCATCTCTCTGAAACAAATTC 1586
Db      564 TGAAGACAGCGGAGAAAGCCCTGCTCTTAAGATTCATCTCTCTGAAACAAATTC 623
Qy      1587 TTTCCACTTACCAAGATGACCTCT 1609
Db      624 TTTCCACTTACCAAGATGACCTCT 646

```

## RESULT 6

BQ431497

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

BQ431497

852 bp

mRNA

linear

EST 24-MAY-2002

AGENCOURT 7894690 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:6158144

5', mRNA sequence.

BQ431497

BQ431497.1 GI:21170583

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;



REFERENCE 1 (bases 1 to 852)  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
 Tissue Procurement: ATCC/DCTD/DRP  
 cDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
 Plate: L1AM13505 row: C column: 09  
 High quality sequence stop: 741.  
 Location/Qualifiers  
 1..852  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6158144"  
 /issue\_type="melanotic melanoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_72"  
 /note="Organ: skin; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."

ORIGIN

Query Match 27.1%; Score 586; DB 5; Length 852;  
 Best Local Similarity 99.8%; Pred. No. 6.3e-310;  
 Matches 636; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1461 AGTAGTACAGGCTGGCTGCTTGGTGTCTTCAAGTCTTCAGCAAACTACATGC 1520  
 |||||||  
 1 AGTAGTACAGGCTGGCTGCTTGGTGTCTTCAAGTCTTCAGCAAACTACATGC 60  
 |||||||  
 1521 ATCCATGAAAGACAGCGGAAAGCCCTGCTCCTAATATCCATCTCTCGAACAAC 1580  
 |||||||  
 61 ATCCATGAAAGACAGCGGAAAGCCCTGCTCCTAATATCCATCTCTCGAACAAC 120  
 |||||||  
 1581 AAATCTTTCCTACTTACCAATGACCCCTCAATCCCATCAATATGATGGTCCAGAAC 1640  
 |||||||  
 121 AAATCTTTCCTACTTACCAATGACCCCTCAATCCCATCAATATGATGGTCCAGAAC 180  
 |||||||  
 1641 CGGATAGCCCCCTTTTATGGTCTCTACAACTAGAGAGAACTCCAAAGAACACACC 1700  
 |||||||  
 181 CGGATAGCCCCCTTTTATGGTCTCTACAACTAGAGAGAACTCCAAAGAACACACC 240  
 |||||||  
 1701 AGATGAAATTTTGGAGCAATGTGTTGTTTTTGGCTGAGCATTAAGATAGGATTA 1760  
 |||||||  
 241 AGATGAAATTTTGGAGCAATGTGTTGTTTTTGGCTGAGCATTAAGATAGGATTA 300  
 |||||||  
 1761 TCTATTCAGAAAGAGCTCAGACATTTCTTAACTAGATGGATCTTAATAGT 1820  
 |||||||  
 301 TCTATTCAGAAAGAGCTCAGACATTTCTTAACTAGATGGATCTTAATAGT 360  
 |||||||  
 1821 TTGCTTCTCAAGAGATGCTCTGTGTTGGGAGAGAGAGCCCAAGAAATATGACAGA 1880  
 |||||||  
 361 TTGCTTCTCAAGAGATGCTCTGTGTTGGGAGAGAGAGCCCAAGAAATATGACAGA 420  
 |||||||  
 1881 CAACATCAGCTTATGAGCCAGAGTGGCGAGATCTCTCTCAGAGAAAGCCCATAT 1940  
 |||||||  
 421 CAACATCAGCTTATGAGCCAGAGTGGCGAGATCTCTCTCAGAGAAAGCCCATAT 480  
 |||||||  
 1941 TTATGTGTGTGAGATGCAAGAAATATGGCCAGAGATGATCATGATCCCTTGTGCAAT 2000  
 |||||||  
 481 TTATGTGTGTGAGATGCAAGAAATATGGCCAGAGATGATCATGATCCCTTGTGCAAT 540  
 |||||||  
 2001 AATAGCAAGAGTTGAGTGTGAAAACTAGAGCATGAAAACTCTGGCATTTTAA 2060  
 |||||||

Db 541 AATAGCAAGAGTTGAGTGTGAAAACTAGAGCATGAAAACTCTGGCATTTAA 600  
 |||||||  
 Qy 2061 AGAGAGAAACCGTACTCTTACAGATATTTGGTCATTA 2097  
 |||||||  
 Db 601 AGAGAGAAACCGTACTCTTACAGATATTTGGTCATTA 637  
 |||||||

RESULT 7  
 BI772430 826 bp mRNA linear EST 25-SBP-2001  
 LOCUS 60305576F1 NIH\_MGC\_122 Homo sapiens cDNA clone IMAGE:5205285 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BI772430  
 VERSION BI772430.1 GI:15764008  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 826)  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
 Plate: L1AM1514 row: 1 column: 22  
 High quality sequence stop: 824.  
 Location/Qualifiers  
 1..826  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5205285"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH\_MGC\_122"  
 /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is Oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 27.1%; Score 568; DB 4; Length 826;  
 Best Local Similarity 99.8%; Pred. No. 5e-300;  
 Matches 618; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGAGAGGTTTCTGTTACTATATGCTACACAGAGGAGCAAGGCGATCGAGAA 60  
 |||||||  
 53 ATGAGAGGTTTCTGTTACTATATGCTACACAGAGGAGCAAGGCGATCGAGAA 112  
 |||||||  
 61 GAAATATGTGAGCAAGCTGTGTACATGATTTTCTGACAGATCTTCACTGATTAAGTAA 120  
 |||||||  
 113 GAAATATGTGAGCAAGCTGTGTACATGATTTTCTGACAGATCTTCACTGATTAAGTAA 172  
 |||||||  
 121 TCCGATATGATTAACCTAAACCGGAAACAGCTCTTGTGTGTGTGTTTACACAG 180  
 |||||||  
 173 TCCGATATGATTAACCTAAACCGGAAACAGCTCTTGTGTGTGTGTTTACACAG 232  
 |||||||  
 181 GGCACCGGAGACCCACCGGACAGAGCCGCAAGTTGTTAAGAAATACAGAAACAAACA 240  
 |||||||  
 233 GGCACCGGAGACCCACCGGACAGAGCCGCAAGTTGTTAAGAAATACAGAAACAAACA 292  
 |||||||

QY 241 CTGCCGTTGATTTCTTTCTCACCCTGCGTATGAGTTACTGGGTTCTCGGTATTCAGAA 300  
 DB 293 CTGCCGTTGATTTCTTTCTCACCCTGCGTATGAGTTACTGGGTTCTCGGTATTCAGAA 352  
 QY 301 TACACCTACTTTTGGCAATGGGGGAAAGATTAATTGATTAACGACTTCAGAGCTTGGAGCC 360  
 DB 353 TACACCTACTTTTGGCAATGGGGGAAAGATTAATTGATTAACGACTTCAGAGCTTGGAGCC 412  
 QY 361 CGGCAATTTCTATGACATCGGACATGCAATGACTGTGTAGTTTAACTTTGGTTGAG 420  
 DB 413 CGGCAATTTCTATGACATCGGACATGCAATGACTGTGTAGTTTAACTTTGGTTGAG 472  
 QY 421 CGGTGATTTGCTGAGACTCTGGCAGAGCCCTCAGAAACATTTTAGTCAAGAGAGACAA 480  
 DB 473 CGGTGATTTGCTGAGACTCTGGCAGAGCCCTCAGAAACATTTTAGTCAAGAGAGACAA 532  
 QY 481 GAGAGATTAATGAGCGCACTCCCGGTGAGCATCCTGCATCTTGAAGACAGACTTGTG 540  
 DB 513 GAGAGATTAATGAGCGCACTCCCGGTGAGCATCCTGCATCTTGAAGACAGACTTGTG 592  
 QY 541 AAGTCAGAGCTGCTACACATTTGATCTCAAGTCGAGCTTCTGAGATTGCATGATTGAGA 600  
 DB 593 AAGTCAGAGCTGCTACACATTTGATCTCAAGTCGAGCTTCTGAGATTGCATGATTGAGA 652  
 QY 601 AGAAGGATTTGAGGTTT 619  
 DB 653 AGAAGGATTTGAGGTTT 671

# RESULT 8

LOCUS AU279788 565 bp mRNA linear EST 31-JUL-2003  
 DEFINITION AU279788 CHONS2 Homo sapiens cDNA clone CHONS2001448 5', mRNA  
 sequence.

ACCESSION AU279788  
 VERSION AU279788.1 GI:28299015  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCES  
 AUTHORS Bunkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 TITLE Imabayashi, H., Mori, T., Gojo, S., Kiyono, T., Sugiyama, T., Irie, R.,  
 Isogai, T., Hata, J., Tomoya, Y., and Umezawa, A.  
 Redifferentiation of dedifferentiated chondrocytes and  
 chondrogenesis of human bone marrow stromal cells via chondrosphere  
 formation with expression profiling by large-scale cDNA analysis  
 JOURNAL Exp. Cell Res. 288 (1), 35-50 (2003)

COMMENT  
 JOURNAL MEDLINE  
 PUBMED 12878157  
 CONTACT: Takao Isogai  
 Genomics Laboratory  
 Helix Research Institute  
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
 Tel: 81-438-52-3975  
 Fax: 81-438-52-3986  
 Email: genomics@hri.co.jp

HRI human cDNA project, Sugiyama, T.; Wakamatsu, A.; Irie, R.;  
 Umezawa, A.; Fukuma, M.; Kusakari, S.; Hata, J.; Ishii, S.; Yamamoto, J.;  
 Iseno, Y.; Saito, K.; Nakamura, Y.; Masuno, Y.; Nagai, K.; Isogai, T.  
 HRI human cDNA project, cDNA library construction & 5'-end one  
 pass sequencing; Helix Research Institute.  
 Location/Qualifiers

FEATURES  
 source  
 1. 565  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CHONS2001448"  
 /cell\_type="Chondrocytes"  
 /clone\_lib="CHONS2"  
 /note="Vector: pME18SFL3"

ORIGIN

Query Match 26.9%; Score 565; DB 1; Length 565;  
 Best Local Similarity 100.0%; Pred. No. 2,1e-298;  
 Matches 565; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 819 TTTTCAAGTCCCAATTTTCAAGGCAAGTCACTTCACTACGAATGATGCCATTAACAC 878  
 DB 1 TTTTCAAGTCCCAATTTTCAAGGCAAGTCACTTCACTACGAATGATGCCATTAACAC 60  
 QY 879 TCTGCTGAGTAAGATTGGAATTTCCTTCAATTCAGACTTTTCTATCAGCTTGAATGATGCC 938  
 DB 61 TCTGCTGAGTAAGATTGGAATTTCCTTCAATTCAGACTTTTCTATCAGCTTGAATGATGCC 120  
 QY 939 CAGCGTATCTGCTTCAAGTATTTGAGGTGATCAAAAGCTTACTTCAAAAGCTGAGCT 998  
 DB 121 CAGCGTATCTGCTTCAAGTATTTGAGGTGATCAAAAGCTTACTTCAAAAGCTGAGCT 180  
 QY 999 TGAAGATTAAGAGAGCACTGCGCTCTTTGAAATTAAGGACACACAAAGAGAAAG 1058  
 DB 181 TGAAGATTAAGAGAGCACTGCGCTCTTTGAAATTAAGGACACACAAAGAGAAAG 240  
 QY 1059 AGCTACCTTACCCCAAGCATATACCTGCGGAGATTTCTTCCAGTTCATTTTACCTGAG 1118  
 DB 241 AGCTACCTTACCCCAAGCATATACCTGCGGAGATTTCTTCCAGTTCATTTTACCTGAG 300  
 QY 1119 TCTTGAATTCGAGCAATCTTAAAGGCAATTTTTCGAGGCTTGTGACTATACAG 1178  
 DB 301 TCTTGAATTCGAGCAATCTTAAAGGCAATTTTTCGAGGCTTGTGACTATACAG 360  
 QY 1179 TGACAGTGTGTAAGGCGGAGGCTACAGAGCTGTGACAGTAAACAAGGGGAGCCGATTA 1238  
 DB 361 TGACAGTGTGTAAGGCGGAGGCTACAGAGCTGTGACAGTAAACAAGGGGAGCCGATTA 420  
 QY 1239 TACCGCTTTGTACAGAGATGCTGTGCTTGTGATCTCTCTCTGCTTCCCTTC 1298  
 DB 421 TACCGCTTTGTACAGAGATGCTGTGCTTGTGATCTCTCTCTGCTTCCCTTC 480  
 QY 1299 TTGCCAGCCACCACTCAGTCTCTGCTGCAACATCTTCTTAACCTCAACCCAGACATA 1358  
 DB 481 TTGCCAGCCACCACTCAGTCTCTGCTGCAACATCTTCTTAACCTCAACCCAGACATA 540  
 QY 1359 TTGCTGTGCAAGCTCAAGTTTATTT 1383  
 DB 541 TTGCTGTGCAAGCTCAAGTTTATTT 565

RESULT 9  
 LOCUS AU124440 877 bp mRNA linear EST 01-AUG-2002  
 DEFINITION AU124440 NT2RM4 Homo sapiens cDNA clone NT2RM4000010 5', mRNA  
 sequence.  
 ACCESSION AU124440  
 VERSION AU124440.1 GI:10949156  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCES  
 AUTHORS Bunkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 TITLE 1 (bases 1 to 877)  
 Ota, T., Wakamatsu, A., Ozawa, M., Ishii, S., Saito, K., Yamamoto, J.,  
 Nakamura, Y., Nishikawa, T., Nagai, T., Suzuki, Y., Sugano, S., and  
 Isogai, T.  
 HRI human cDNA project (Ota, T., Wakamatsu, A., Ozawa, M., Ishii, S.,  
 Saito, K., Yamamoto, J., Nakamura, Y., Nishikawa, T., Nagai, T.,  
 Suzuki, Y., Sugano, S., Isogai, T.)  
 Unpublished (2000)

COMMENT  
 JOURNAL MEDLINE  
 PUBMED 12878157  
 CONTACT: Takao Isogai  
 Genomics Laboratory  
 Helix Research Institute  
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
 Tel: 81-438-52-3975  
 Fax: 81-438-52-3986  
 Email: genomics@hri.co.jp

HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix Research Institute; cDNA library construction; Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

## FEATURES

source

location/Qualifiers

1..877  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="NT2RM400010"  
/cell\_type="teratocarcinoma"  
/cell\_line="NT2"  
/clone\_id="NT2RM4"  
/notes="Vector: DME18SFu3; mRNA from uninduced NT2 neuronal precursor cells"

## ORIGIN

Query Match 25.9%; Score 543; DB 1; Length 877;  
Best Local Similarity 100.0%; Pred. No. 2.9e-286;  
Matches 543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1143 AAGGCAATTTTGAGAGCCCTTGTGACTATACAGTGAAGTCTGAAGGCGAGGCT 1202  
192 AAGGCAATTTTGAGAGCCCTTGTGACTATACAGTGAAGTCTGAAGGCGAGGCT 251  
1203 ACAGAGCTTGAGTGAAGAACAGGGGCGAGCGATTATAGCCGTTGTAGAGATGCTG 1262  
252 ACAGAGCTTGAGTGAAGAACAGGGGCGAGCGATTATAGCCGTTGTAGAGATGCTG 311  
1263 TGGCTGTTGTGATCTCTCTGCTTCCCTTCCCTTCTTCCAGCCACCATCTGCTCT 1322  
312 TGGCTGTTGTGATCTCTCTGCTTCCCTTCCCTTCTTCCAGCCACCATCTGCTCT 371  
1323 GCTGGAACATCTTCTTAACTTCAACCCAGACCATATTCGTCGCAAGCTCAAGTTATT 1382  
372 GCTGGAACATCTTCTTAACTTCAACCCAGACCATATTCGTCGCAAGCTCAAGTTATT 431  
1383 TCACCCAGAAAGCTCCATTTTGTCTTCAACATTTGGAATTTCTGTCTACGCAAC 1442  
432 TCACCCAGAAAGCTCCATTTTGTCTTCAACATTTGGAATTTCTGTCTACGCAAC 491  
1443 AGAGGTTCTGCGAAGGAGTATGTACAGGCTGGCTGTGTGTTGCTTCAAGTTCT 1502  
492 AGAGGTTCTGCGAAGGAGTATGTACAGGCTGGCTGTGTGTTGCTTCAAGTTCT 551  
1503 TCAGCCAAACATACATCATCCCATGAAGAAGGCGGAAAGCCCTGCTCTAAGATATC 1562  
552 TCAGCCAAACATACATCATCCCATGAAGAAGGCGGAAAGCCCTGCTCTAAGATATC 611  
1563 CATCTCTCTCGAACCAAAATCTTTCACATTAACAGATGACCCCTCAATCCCATCAT 1622  
612 CATCTCTCTCGAACCAAAATCTTTCACATTAACAGATGACCCCTCAATCCCATCAT 671  
1623 AATGTGGGTTCAGGAACCGGCAATAGCCGTTTATTTGGTCTCTAACAATAGAGAA 1682  
672 AATGTGGGTTCAGGAACCGGCAATAGCCGTTTATTTGGTCTCTAACAATAGAGAA 731  
1683 ACT 1685  
732 ACT 734

RESULT 10  
BQ218755 1061 bp mRNA linear EST 02-MAY-2002  
LOCUS BQ218755  
DEFINITION BQ218755 1061 bp mRNA linear EST 02-MAY-2002  
5', mRNA sequence.  
ACCESSION BQ218755  
VERSION BQ218755.1 GI:20400155  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 1061)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt BioScience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: LLM13279 row: n column: 07  
High quality sequence stop: 518.

## FEATURES

source

location/Qualifiers

1..1061  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6041670"  
/issue\_type="embryonal carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_id="NIH\_MGC\_92"  
/note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC library."

## ORIGIN

Query Match 25.3%; Score 531; DB 5; Length 1061;  
Best Local Similarity 100.0%; Pred. No. 1.2e-279;  
Matches 531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

995 AGCTTGAAGTAAAG 1054  
1 AGCTTGAAGTAAAG 60  
1055 AAGGAGCTACCTTACCCAGCATATACCTGCGGAGATGCTCTCCAGTTCATTTTAACT 1114  
61 AAGGAGCTACCTTACCCAGCATATACCTGCGGAGATGCTCTCCAGTTCATTTTAACT 120  
1115 GGTGCTTGAATCCAGCAATCTTAAAGGAGATTTTTCGAGGCCCTTGTGACTATA 1174  
121 GGTGCTTGAATCCAGCAATCTTAAAGGAGATTTTTCGAGGCCCTTGTGACTATA 180  
1175 CCAGTACAGTGTGTAAGGCGAGGCTACAGAGCTGTGCACTAAACAAAGGGGAGCCG 1234  
181 CCAGTACAGTGTGTAAGGCGAGGCTACAGAGCTGTGCACTAAACAAAGGGGAGCCG 240  
1235 AATTATGCGGCTTGTAGAGATGCGTGTGCTGTTGTGATCTCTCTGCTCTGCTTTC 1294  
241 AATTATGCGGCTTGTAGAGATGCGTGTGCTGTTGTGATCTCTCTGCTCTGCTTTC 300  
1295 CTTCTTGCAAGCAGCAGCTCAGTCTCTGCTGCAACATCTTCTAACTTCAACCCAGAC 1354  
301 CTTCTTGCAAGCAGCAGCTCAGTCTCTGCTGCAACATCTTCTAACTTCAACCCAGAC 360  
1355 CATATTCGTGTGCAAGCTCAAGTTATTTTCAACCCAGAAAGCTTCATTTGTCTTCAACA 1414  
361 CATATTCGTGTGCAAGCTCAAGTTATTTTCAACCCAGAAAGCTTCATTTGTCTTCAACA 420  
1415 TTGTGGAATTTCTGTACTAGCCCAACAGAGTTCTGCGGAAGGAGATGTACAGGCT 1474  
421 TTGTGGAATTTCTGTACTAGCCCAACAGAGTTCTGCGGAAGGAGATGTACAGGCT 480  
1475 GGCTGGCCTTGTGTTGCTTCAAGTTCTTCAAGCAACATACATGATGCC 1525  
481 GGCTGGCCTTGTGTTGCTTCAAGTTCTTCAAGCAACATACATGATGCC 531

```

RESULT 11
LOCUS      BU941078
DEFINITION
ACCESSION  BU941078
VERSION    BU941078
KEYWORDS   BU941078.1  GI:24129897
SOURCE     EST.
ORGANISM   Homo sapiens (human)
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 834)
            NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: NCI
            CDNA Library Preparation: Michael Brownstein Laboratory
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNLN at:
            http://image.llnl.gov
            Plate: L12CM3022 row: e column: 21
            High quality sequence stop: 506.
            Location/Qualifiers
            1..834

```

Db	183	ATTGAATCTCAAGTCGAGCTTGTGANTGTGATTCAGGAGAAAGAAATTCGAGGTT	242
Qy	619	TTGAGCAAAATGCAGTGGAACAGCAACCAATCCAATGTTGTAATTTGAAGCTTTGAGTCC	678
Db	243	TTGAAGCAAAATGCAGTGGAACAGCAACCAATCCAATGTTGTAATTTGAAGACTTTGAGATCC	302
Qy	679	TCACCTTACCCGTTGGGTACCCCACTCTCACAAAGCTCTCTGAATATTCTCTGGTTTACCC	738
Db	303	TCACCTTACCCGTTGGGTACCCCACTCTCACAAAGCTCTCTGAATATTCTCTGGTTTACCC	362
Qy	739	CCAGAAATTTTACAGGTACATCTGCAGAGAAGTCTTTGGCCAGAGAGAAAGCCAACTATCT	798
Db	363	CCAGAAATATTTACAGGTACATCTGCAGAGAAGTCTTTGGCCAGAGAGAAAGCCAACTATCT	422
Qy	799	GTGACTTCAGAGATTCAGTCTTTCGAAGTGCATTTCAAGGAGAGTTCAACTTAACG	858
Db	423	GTGACTTCAGAGATTCAGTCTTTCGAAGTGCATTTCAAGGAGAGTTCAACTTAACG	482
Qy	859	AATGATGCAATAAAAAACAATCTGCTGCTGTAATTTGCAATTTCAAAATPACAGACTTTTCC	918
Db	483	AATGATGCAATAAAAAACAATCTGCTGCTGTAATTTGCAATTTCAAAATPACAGACTTTTCC	542
Qy	919	TATGAGCTGAGAGATGCTTCAGGCTATCTGCGCTTAACAGTGAATTCTGAGGTCAAAAGC	978
Db	543	TATGAGCTGAGAGATGCTTCAGGCTATCTGCGCTTAACAGTGAATTCTGAGGTCAAAAGC	602
Qy	979	CTACTCCAA 987	
Db	603	CTACTCCAA 611	

```

/db_xref="taxon:9606"
/c1one="IMAGE:6712893"
/tissue_type="mixed (pool of 40 RNAs)"
/lab_host="DH10B (T1-phage-resistant)"
/c1one_id="NIH_MGC_128"
/note="Vector: pDNR-LIB; Site 1: SfiI (ggccatcatggcc);
Site 2: SfiI (ggccgcctggcc); Double-stranded cDNA was
prepared from a pool of 40 cell line polyA+ RNAs (bladdered
- 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon
4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,
kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell -
5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%,
salivary gland - 1.3%, and skin - 2.3%). 5' and 3'
adaptors were used in cloning as follows:
5'-AAGCAGTGCATTCACCCAGATGGCATTCACGCCCGC-3' and
5'-ATTCTAAGGCCGAGCGCGCGCAGCATG-dT(30)NM-3'. Full-length
enriched library was constructed using the Clontech
Creator SMART kit and size-selected to contain the >2 kb
size fraction (other fractions present in NIH_MGC_126 and
NIH_MGC_127). Library created in the laboratory of T.
Uedán, W.D., Ph.D. (NIH, NIH). Note: this is a NIH_MGC
Library."

```

ORIGIN

Query Match	24.2%	Score 507	DB 5	Length 834
Best Local Similarity	99.7%	Pred. No. 1.8e-266		
Matches 607; Conservative	0	Mismatches 2	Indels 0	Gaps 0

[illegible]

RESULT	12
LOCUS	CB164340
DEFINITION	CB164340 521 bp mRNA linear EST 30-JAN-2003 K-EST0225498 L17N670205n1 Homo sapiens cDNA clone L17N670205n1-39-F02 5', mRNA sequence.
ACCESSION	CB164340
VERSION	CB164340.1
KEYWORDS	GI:28150466
SOURCE	EST.
ORGANISM	Homo sapiens (human)
TITLE	Homo sapiens
JOURNAL	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 521) Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.
COMMENT	21C Frontier Korean Est Project 2001 Unpublished (2002) Contact: Kim YS Genome Research Center Korea Research Institute of Bioscience & Biotechnology 52 Boseun-dong Yuseong-gu, Daejeon 305-333, South Korea Tel.: +82-42-860-4470 Fax: +82-42-860-4409 Email: yongsang@mail.kribb.re.kr Plate: 39 row: F column: 02 High quality sequence, stop: 521. Location/Qualifiers
FEATURES	

```

source
1. .521
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="L17N670205n1-39-F02"
/bex="F"
/lab_host="Top10F"
/clone_id="L17N670205n1"
/notes="Organ: Liver; Vector: pT7T3-Pac; Site 1: EcoRI;
Site 2: NotI; The library was constructed by the Soares
laboratory and it was constructed as described by Bonaldo,
M.F., Lemmon, G. and Soares, M.B. (1996), Genome Research
6(9): 791-806. RNA was prepared from harvested cell
culture."

```

/nucleo-Organ: liver; Vector: pT73-Pac; Site\_1: EcoRI; Site\_2: NotI; The library was contributed by the Soares Laboratory and it was constructed as described by Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome Research 6(9): 791-806. RNA was prepared from harvested cell culture."

ORIGIN

Query Match 22.4%; Score 470; DB 6; Length 521;  
Best Local Similarity 99.8%; Pred. No. 3.9e-246;  
Matches 520; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1540 AAGCCCTGCTCTTAAGATATTCATCTCTCTGAAACAATTTTCCACTTACCA 1599  
DB 1 AAGGCCCTGCTCTTAAGATATTCATCTCTCTGAAACAATTTTCCACTTACCA 60  
QY 1600 GATGACCCCTCAATCCCATCATTAATGTGGTCCAGAAACCGCATAGCCCGTTTATT 1659  
DB 61 GATGACCCCTCAATCCCATCATTAATGTGGTCCAGAAACCGCATAGCCCGTTTATT 120  
QY 1660 GGGTTCCTACACATAGAGAGAACTCCAAAGAACACACCCAGATGGAATTTTGGAGCA 1719  
DB 121 GGGTTCCTACACATAGAGAGAACTCCAAAGAACACACCCAGATGGAATTTTGGAGCA 180  
QY 1720 ATGTGCTGTTTGTGCTGCTGAGCAATAGATAGGATTTATCTATTCAGAAAAGAGCTC 1779  
DB 181 ATGTGCTGTTTGTGCTGCTGAGCAATAGATAGGATTTATCTATTCAGAAAAGAGCTC 240  
QY 1780 AGACATTTCTTAAGCATGAGGATCTTAATCATCTAAAGGTTTCTTCTCAAGAGATGCT 1839  
DB 241 AGATATTTCTTAAGCATGAGGATCTTAATCATCTAAAGGTTTCTTCTCAAGAGATGCT 300  
QY 1840 CCGTGTGGGAGAGAGAAAGCCCGCAGAAATATGTACAGAACATCCAGCTTCATGCG 1899  
DB 301 CCGTGTGGGAGAGAGAAAGCCCGCAGAAATATGTACAGAACATCCAGCTTCATGCG 360  
QY 1900 CACACAGTGGCGAATATCTCTCCAGAGAAACGCGCATTTTATGTGTGTGAGAGATGCA 1959  
DB 361 CACACAGTGGCGAATATCTCTCCAGAGAAACGCGCATTTTATGTGTGTGAGAGATGCA 420  
QY 1960 AAGAATATGCGCAGATGTATCATATGATGCTTGTGTGCAATTAATTAAGCAAAAGATTGA 2019  
DB 421 AAGAATATGCGCAGATGTATCATATGATGCTTGTGTGCAATTAATTAAGCAAAAGATTGA 480  
QY 2020 GTTGAATACTAAGACATGAAAACTCTGCGCATTTTAA 2060  
DB 481 GTTGAATACTAAGACATGAAAACTCTGCGCATTTTAA 521

RESULT 13  
CB97527 776 bp mRNA linear EST 01-MAY-2003  
LOCUS AGENCOURT13620640 NIH MGC 148 Homo sapiens cDNA clone  
DEFINITION IMAGE:30338684 5', mRNA sequence.  
CB97527  
VERSION CB97527.1 GI:30292047  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Bukayrota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euteleia; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 776)  
NIH-MGC <http://mgs.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strauberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: Dr. Stefan Hanson  
CDNA Library Preparation: Michael J. Brownstein (NHGRI) with help  
and advice from Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.lnl.gov>  
Plate: NDAM365 row: 1 column: 21  
High quality sequence stop: 564.  
Location/Qualifiers  
1..776

FEATURES  
source

ORIGIN

Query Match 22.0%; Score 461; DB 6; Length 776;  
Best Local Similarity 99.5%; Pred. No. 3.7e-241;  
Matches 611; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGAGAGGTTCTGTACTATATGCTACACAGCAGGACAGCAAGCCATCGCAGAA 60  
DB 88 ATGAGAGGTTCTGTACTATATGCTACACAGCAGGACAGCAAGCCATCGCAGAA 147  
QY 61 GAAATATGTAGAGCAGCTGTGTACATGATTTTCTGACAGATCTTCACTGTATTAGTAA 120  
DB 148 GAAATGTGTAGAGCAGCTGTGTACATGATTTTCTGACAGATCTTCACTGTATTAGTAA 207  
QY 121 TCCGATTAATATGCTTAAACCCGAAACAGCTCTCTTGTGTGTGTTTCTACACG 180  
DB 208 TCCGATTAATATGCTTAAACCCGAAACAGCTCTCTTGTGTGTGTTTCTACACG 267  
QY 181 GGCACCGGAGACCCACCCGACACAGCCCGCAAGTTTGTAAAGAAATACGAACCAACA 240  
DB 268 GGCACCGGAGACCCACCCGACACAGCCCGCAAGTTTGTAAAGAAATACGAACCAACA 327  
QY 241 CTGCGGTTGATTTCTTGTCTCACTGCGGTATGAGTTA CTGCGTCTCGGTATTCAGAA 300  
DB 328 CTGCGGTTGATTTCTTGTCTCACTGCGGTATGAGTTA CTGCGTCTCGGTATTCAGAA 387  
QY 301 TACACCTACTTTTGCATATGCGGAGAAATATTTGATTAACGACTTCAAGGCTTGAAGC 360  
DB 388 TACACCTACTTTTGCATATGCGGAGAAATATTTGATTAACGACTTCAAGGCTTGAAGC 447  
QY 361 CGGCATTTCTATGACCTGACATGACATGACATGACATGACATGACATGACATGACATGAC 420  
DB 448 CGGCATTTCTATGACCTGACATGACATGACATGACATGACATGACATGACATGACATGAC 507  
QY 421 CCGTGATTTCTGATCTGCGCCAGCCCTCAGAAACATTTTAGTCAAGCAGAGCA 480  
DB 508 CCGTGATTTCTGATCTGCGCCAGCCCTCAGAAACATTTTANGTCAAGCAGAGCA 567  
QY 481 GAGAGATTAATGTGGCAGCTCCCGGTGGCATCTGTCATCTTGAAGACAGACCTTGTG 540  
DB 568 GAGAGATTAATGTGGCAGCTCCCGGTGGCATCTGTCATCTTGAAGACAGACCTTGTG 627  
QY 541 AAGTCAGAGCTGTACATGATTAATCAAGTGAAGCTTGAAGTTCATGATTCAGGA 600  
DB 628 AAGTCAGAGCTGTACATGATTAATCAAGTGAAGCTTGAAGTTCATGATTCAGGA 687  
QY 601 AGAAGGATTTCTGA 614  
DB 688 AGAAGGATTTCTGA 701

RESULT 14  
AU132586 822 bp mRNA linear EST 01-AUG-2002  
LOCUS AU132586 NT2RPA Homo sapiens cDNA clone NT2RPA000141 5', mRNA  
DEFINITION sequence.

Accession	Version	KeyWords	Source	Organism	Reference	Authors	Title	Journal	Comment	Features	Source	Origin
AUI32586	AUI32586.1	GI:10992940	EST.	Homo sapiens (human)	1 (bases 1 to 822)	Ota, T., Sugiyama, T., Ishii, S., Suzuki, Y., Saito, K., Yamamoto, J., Nishikawa, T., Nakamura, Y., Nagai, T., Sugano, S., Masuno, Y. and Isogai, T.	HRI human cDNA project (Ota, T., Sugiyama, T., Ishii, S., Suzuki, Y., Saito, K., Yamamoto, J., Nishikawa, T., Nakamura, Y., Nagai, T., Sugano, S., Masuno, Y., Isogai, T.)	Unpublished (2000)	Contact: Takao Isogai Genomics Laboratory Helix Research Institute 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan Tel.: 81-438-52-3975 Fax: 81-438-52-3986 Email: genomics@hri.co.jp HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.	Location/Qualifiers 1..822	Query Match Best Local Similarity 99.6%; Pred. No. 7.4e-218; Matches 555; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
Db	182	GGTCTCGGTGATTCAGAAATACACCTACTTTTGCATGCGGGGAAAGATATGATTAACGA	283	GGTCTCGGTGATTCAGAAATACACCTACTTTTGCATGCGGGGAAAGATATGATTAACGA	342							
Db	182	GGTCTCGGTGATTCAGAAATACACCTACTTTTGCATGCGGGGAAAGATATGATTAACGA	241	GGTCTCGGTGATTCAGAAATACACCTACTTTTGCATGCGGGGAAAGATATGATTAACGA	241							
Qy	343	CTTCAAGCTTGAAGCCCGGCATTTCTATGACACTGGAACATGCAGATGACTGTGTAGT	402	CTTCAAGCTTGAAGCCCGGCATTTCTATGACACTGGAACATGCAGATGACTGTGTAGT	402							
Db	242	CTTCAAGCTTGAAGCCCGGCATTTCTATGACACTGGAACATGCAGATGACTGTGTAGT	301	CTTCAAGCTTGAAGCCCGGCATTTCTATGACACTGGAACATGCAGATGACTGTGTAGT	301							
Qy	403	TTAGAACCTTGCTGTTAGCCGCTGATGCTGGAACCTGCGCCAGCCCTGAGAAAGCATTTT	462	TTAGAACCTTGCTGTTAGCCGCTGATGCTGGAACCTGCGCCAGCCCTGAGAAAGCATTTT	462							
Db	302	TTAGAACCTTGCTGTTAGCCGCTGATGCTGGAACCTGCGCCAGCCCTGAGAAAGCATTTT	361	TTAGAACCTTGCTGTTAGCCGCTGATGCTGGAACCTGCGCCAGCCCTGAGAAAGCATTTT	361							
Qy	463	AGGTCAAGCAGAGGACAGAGAGATTAAGTGGCGCACTCCCGTGGCATCACTGCATCC	522	AGGTCAAGCAGAGGACAGAGAGATTAAGTGGCGCACTCCCGTGGCATCACTGCATCC	522							
Db	362	AGGTCAAGCAGAGGACAGAGAGATTAAGTGGCGCACTCCCGTGGCATCACTGCATCC	421	AGGTCAAGCAGAGGACAGAGAGATTAAGTGGCGCACTCCCGTGGCATCACTGCATCC	421							
Qy	523	TTGAGGACAGACCTTTGTGAAGTCAAGAGTCTCAACATTGAATCTCAAGTGAAGCTTCTG	582	TTGAGGACAGACCTTTGTGAAGTCAAGAGTCTCAACATTGAATCTCAAGTGAAGCTTCTG	582							
Db	422	TTGAGGACAGACCTTTGTGAAGTCAAGAGTCTCAACATTGAATCTCAAGTGAAGCTTCTG	481	TTGAGGACAGACCTTTGTGAAGTCAAGAGTCTCAACATTGAATCTCAAGTGAAGCTTCTG	481							
Qy	583	AGATTTCAGATTCACAGGAAGAAAGATTCGAGGTTTGAAGCAAAATGACAGTGAACGC	642	AGATTTCAGATTCACAGGAAGAAAGATTCGAGGTTTGAAGCAAAATGACAGTGAACGC	642							
Db	482	AGATTTCAGATTCACAGGAAGAAAGATTCGAGGTTTGAAGCAAAATGACAGTGAACGC	541	AGATTTCAGATTCACAGGAAGAAAGATTCGAGGTTTGAAGCAAAATGACAGTGAACGC	541							
Qy	643	AACCAATCAATGTTGTATTAAGAACTTTGAGTCTCACTTACCCGTTCCGTACCCCA	702	AACCAATCAATGTTGTATTAAGAACTTTGAGTCTCACTTACCCGTTCCGTACCCCA	702							
Db	542	AACCAATCAATGTTGTATTAAGAACTTTGAGTCTCACTTACCCGTTCCGTACCCCA	601	AACCAATCAATGTTGTATTAAGAACTTTGAGTCTCACTTACCCGTTCCGTACCCCA	601							
Qy	703	CTCTCACAAGCTCTCTGAATATTCCTGCTTACCCCGCAATATTTACAGTACATCTG	762	CTCTCACAAGCTCTCTGAATATTCCTGCTTACCCCGCAATATTTACAGTACATCTG	762							

Db	602	CTCTCACAAGGCTCTGAAATATTCGTGTTACCCCAAGATATTACAGTAGACATCTG	661					
Qy	763	CAGAGATCTCTTGGCCAGAGAGAAAGCAGATATCTGTACTTCAGACAGATCCAGTTTTT	822					
Db	662	CAGAGATCTCTTGGCCAGAGAGAAAGCAGATATCTGTGACTTCAGACAGATCCAGTTTTT	721					
Qy	823	CAAGTGCATTTTCAAA	839					
Db	722	CAGTGCATTTTCAAA	738					
RESULT 15								
LOCUS	AM965709	591 bp	mRNA linear EST 01-JUN-2000					
DEFINITION	EST377782 MAGI resequences, MAGI Homo sapiens cDNA, mRNA sequence.							
ACCESSION	AM965709							
VERSION	AM965709.1	GI:8155545						
KEYWORDS	EST.							
SOURCE	Homo sapiens (human)							
ORGANISM	Homo sapiens							
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.							
	1 (bases 1 to 591)							
	Hedge, P., Qi, R., Abernathy, K., Dharap, S., Gaspar, R., Gay, C.,							
	Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and							
AUTHORS	Quackenbush, J.							
	Assessment of gene expression patterns in a model of colon tumor							
	metastasis using a 19,200 element cDNA microarray							
	Unpublished (2000)							
	Contact: John Quackenbush							
JOURNAL	The Institute for Genomic Research							
	9712 Medical Center Dr., Rockville, MD 20850, USA							
	Tel: 301 838 3528							
	Fax: 301 838 0208							
	Email: johnq@tigr.org							
COMMENT	Plate: 218							
	Seq primer: Reverse.							
	Location/Qualifiers							
	1..591							
	/organism="Homo sapiens"							
FEATURES	/mol_type="mRNA"							
	/db_xref="taxon:9606"							
	/clone_id="MAGI resequences, MAGI"							
	/note="Vector: pBluescriptSk"							
	ORIGIN							
Query Match								
Best Local Similarity 100.0%; Pred. No. Se=23; Length 591;								
Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;								
Qy	1192	AAGCGAGCGCTACAGAGCTGTGCGAGTAACAAGGGGAGCCGATTTAGCCGCTTTGTA	1251					
Db	1	AAGCGAGCGCTACAGAGCTGTGCGAGTAACAAGGGGAGCCGATTTAGCCGCTTTGTA	60					
Qy	1252	CGAGATGCTGTGCTGCTCTTGTGGATCTCTCTCTGCTTCCCTTCTTGGCAGCACCA	1311					
Db	61	CGAGATGCTGTGCTGCTCTTGTGGATCTCTCTCTGCTTCCCTTCTTGGCAGCACCA	120					
Qy	1312	CTCAGTCTCTGCTGCGAACAATCTTCTTAACTTCAACCCAGACATATTGTGTGGACG	1371					
Db	121	CTCAGTCTCTGCTGCGAACAATCTTCTTAACTTCAACCCAGACATATTGTGTGGACG	180					
Qy	1372	TCAAGTTTATTTACCCAGAGAAAGCTCATTTGTCTTCAACATTTGGAATTTCTGCT	1431					
Db	181	TCAAGTTTATTTACCCAGAGAAAGCTCATTTGTCTTCAACATTTGGAATTTCTGCT	240					
Qy	1432	ACTGCGACAACAGAGGTTCTGCGGAAGGAGATATGACAGGCTGAGCTGGCTTGTGGTT	1491					
Db	241	ACTGCGACAACAGAGGTTCTGCGGAAGGAGATATGACAGGCTGAGCTGGCTTGTGGTT	300					
Qy	1492	GCTTCATTTCTTCCAGCCAAATATACATGATCCCATGAAACAACGCGGGAAGCCCTGGCT	1551					
Db	301	GCTTCATTTCTTCCAGCCAAATATACATGATCCCATGAAACAACGCGGGAAGCCCTGGCT	360					

Qy	1552	CCTAAGATATCCATCTCTCTCGAACAATAATTCTTCCACTTACCAATGACCCCTCA	1611
Db	361	CCTAAGATATCCATCTCTCTCGAACAATAATTCTTCCACTTACCAATGACCCCTCA	420
Qy	1612	ATCCCCATCATATGATGGTCCAGGAA	1639
Db	421	ATCCCCATCATATGATGGTCCAGGAA	448

Search completed: August 27, 2005, 15:58:33  
Job time : 4548.04 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 19:17:21 ; Search time 6007.21 Seconds  
(without alignments)  
16914.771 Million cell updates/sec

Title: US-09-371-347A-41

Perfect score: 2097  
Sequence: 1 atgagagaggttcctgctact.....ttcagatattcgtcataa 2097

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_rc:\*  
11: gb\_scs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2046	97.6	3241	6	COJ26091 Sequence
2	2046	97.6	3259	6	ARI44976 Sequence
3	2046	97.6	3259	6	AX050463 Sequence
4	2046	97.6	3259	9	AF025794 Homo sapi
5	2046	97.6	3291	9	AF121214 Homo sapi
6	1995	95.1	3310	9	BC054816 Homo sapi
7	1338	63.8	2933	11	BV177620 Homo sapi
8	1338	63.8	2933	11	BV178010 Homo sapi
9	386	18.4	330	6	BD077780 S. EST of
10	381	18.2	1353	9	AF121205 Homo sapi
11	330	15.7	109626	9	AC010346 Homo sapi
12	330	15.7	110756	9	AC025174 Homo sapi
13	279	13.3	158199	2	AC022921 Homo sapi
14	279	13.3	167237	2	AC021609 Homo sapi
15	279	13.3	177596	2	AC091945 Homo sapi
16	189	9.0	1156	9	AF12120509 Homo sapi
17	188	9.0	1034	9	AF121212 Homo sapi
18	188	9.0	167237	2	AC021609 Homo sapi
19	188	9.0	177596	2	AC091945 Homo sapi

c	20	183	8.7	1432	9	AF121208 Homo sapi
	21	161	7.7	158199	2	AC022921 Homo sapi
	22	158	7.5	1256	9	AF121202S02 Homo sapi
	23	158	7.5	2475	6	AR454615 Sequence
	24	158	7.5	2475	6	AX375651 Sequence
	25	155	7.4	2011	9	AF121202S06 Homo sapi
	26	146	7.0	2214	9	AF121202S12 Homo sapi
	27	125	6.0	969	9	AF121202S05 Homo sapi
	28	121	5.8	1119	9	AF121211 Homo sapi
	29	119	5.7	1200	9	AF121202S10 Homo sapi
	30	78	3.7	4506	9	AF121202S01 Homo sapi
	31	65	3.1	244	6	BD024284 S. EST and
	32	63	3.0	63	6	AX611833 Sequence
	33	60	2.9	60	6	CO539377 Sequence
	34	54	2.6	54	6	AX611839 Sequence
	35	54	2.6	54	6	AX611843 Sequence
c	36	51	2.4	51	6	AX162161 Sequence
	37	48	2.3	48	6	AX611835 Sequence
	38	47	2.3	48	6	AX611841 Sequence
	39	47	2.2	183	6	CO670532 Sequence
	40	44	2.1	650	9	AF121209 Homo sapi
	41	41	2.0	41	6	AX611845 Sequence
	42	38	1.8	38	6	AX611837 Sequence
	43	38	1.8	63	6	AX611834 Sequence
c	44	32	1.5	238720	2	AC095949 Rattus no
	45	32	1.5	271339	2	AC131637 Rattus no

#### ALIGNMENTS

RESULT 1	COJ26091	Sequence 12025 from Patent WO02068579.	3241 bp	DNA	linear	PAT 03-FRB-2004
LOCUS	COJ26091					
DEFINITION	Sequence 12025 from Patent WO02068579.					
ACCESSION	COJ26091					
VERSION	COJ26091.1	GI:42288134				
KEYWORDS						
SOURCE						
ORGANISM	Homo sapiens (human)					
REFERENCE						
AUTHORS	Venter, C.J., Adams, M.C., Li, P.W. and Myers, B.W.					
TITLE	Kite, such as nucleic acid arrays, comprising a majority of humenexons or transcripts, for detecting expression and other uses thereof					
JOURNAL	Patent: WO 02068579-A 12025 06-SEP-2002;					
FEATURES						
source	1..3241					
	/organism="Homo sapiens"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:9606"					
ORIGIN						
Query Match	97.6%;	Score 2046;	DB 6;	Length 3241;		
Best Local Similarity	100.0%;	Pred. No. 0;				
Matches 2096;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;		
QY	1	ATGAGAGAGGTTCTGTTACTATATGTTACACAGCAGGACGACGAAAGCCATCGAGAA	60			
DB	80	ATGAGAGAGGTTCTGTTACTATATGTTACACAGCAGGACGACGAAAGCCATCGAGAA	139			
QY	61	GAATATGTGACCAAGCTGTGTACATGATTTTCTGAGATCTTCACTGATATTAGTAA	120			
DB	140	GAATATGTGACCAAGCTGTGTACATGATTTTCTGAGATCTTCACTGATATTAGTAA	199			
QY	121	TTCGATTAAGTATGACCTTAACCAAGCAAGCTCCCTGTTGTTGTTCTTACACAG	180			
DB	200	TTCGATTAAGTATGACCTTAACCAAGCAAGCTCCCTGTTGTTGTTCTTACACAG	259			
QY	181	GGCACCAGAGACCCAGACAGCCCGCAAGTTGTTAAGGAATATACGAACCAAGCA	240			

Db 260 GGACCGGAGACCCACCGACACAGCCGCAAGTTGTTAAGAAATATACGAAACCAACA 319  
 Qy 241 CTCGCGGTGATTTCTTGTCTCACTGCGGTATGGTTATCGGGTCTCGGTGATTCAGAA 300  
 Db 320 CTCGCGGTGATTTCTTGTCTCACTGCGGTATGGTTATCGGGTCTCGGTGATTCAGAA 379  
 Qy 301 TACACCTACTTTTGCAATGGGAGAGATTAATGATTAACGACTTCAAGAGCTTGAAGCC 360  
 Db 380 TACACCTACTTTTGCAATGGGAGAGATTAATGATTAACGACTTCAAGAGCTTGAAGCC 439  
 Qy 361 CGGCAATTTCTATGACACTGACATGACATGACATGACATGACATGACATGACATGACAT 420  
 Db 440 CGGCAATTTCTATGACACTGACATGACATGACATGACATGACATGACATGACATGACAT 499  
 Qy 421 CGGTGATTTCTGGAATCTGCGGACGCTTCAAGAAAGATTTAGTCAAGCAGAGACAA 480  
 Db 500 CGGTGATTTCTGGAATCTGCGGACGCTTCAAGAAAGATTTAGTCAAGCAGAGACAA 559  
 Qy 481 GAGAGATTAAGTGGCGCATCTCCGGTGGCATCACCTGATCTTTGAGAGACAGCTTTGG 540  
 Db 560 GAGAGATTAAGTGGCGCATCTCCGGTGGCATCACCTGATCTTTGAGAGACAGCTTTGG 619  
 Qy 541 AAGTCAGAGCTGCTACATTTGAAATCTCAAGTGGAGCTTCAAGATTCGATTCAGAG 600  
 Db 620 AAGTCAGAGCTGCTACATTTGAAATCTCAAGTGGAGCTTCAAGATTCGATTCAGAG 679  
 Qy 601 AGAAGATTTCTGAGTTTGAAGCAAAATGACGTGAACAGCAACCAATGTTGTA 660  
 Db 680 AGAAGATTTCTGAGTTTGAAGCAAAATGACGTGAACAGCAACCAATGTTGTA 739  
 Qy 661 ATTGAAGCTTTGAGTCTCTCACTTCCGTTGGATCCCGCATCTCTCAAGCCTCTCTG 720  
 Db 740 ATTGAAGCTTTGAGTCTCTCACTTCCGTTGGATCCCGCATCTCTCAAGCCTCTCTG 799  
 Qy 721 AATATCTGTTTACCCCGAAGATTTTACAGGTATCATGCGAGAGCTCTTGGGCGAG 780  
 Db 800 AATATCTGTTTACCCCGAAGATTTTACAGGTATCATGCGAGAGCTCTTGGGCGAG 859  
 Qy 781 GAGAAAGCCAAAGTATCTGATCTTCAAGAGATCAGTTTTCAGTGGCAATTTCAAG 840  
 Db 860 GAGAAAGCCAAAGTATCTGATCTTCAAGAGATCAGTTTTCAGTGGCAATTTCAAG 919  
 Qy 841 GCAATCACTTACTGATGATGATGCAATTAACCACTGCTGCTGATTAATGATGAT 900  
 Db 920 GCAATCACTTACTGATGATGATGCAATTAACCACTGCTGCTGATTAATGATGAT 979  
 Qy 901 TCAATCACTTACTGATGATGATGCAATTAACCACTGCTGCTGATTAATGATGAT 960  
 Db 980 TCAATCACTTACTGATGATGATGCAATTAACCACTGCTGCTGATTAATGATGAT 1039  
 Qy 961 GATTTGAGGTACAAAGCTTCTCAAGAGCTGAGCTTGAAGATTAAGAGAGAGAGCTGC 1020  
 Db 1040 GATTTGAGGTACAAAGCTTCTCAAGAGCTGAGCTTGAAGATTAAGAGAGAGAGAGCTGC 1099  
 Qy 1021 GTCCTTTGAAATTAAGGAGAGACACAAAGAAAGAGAGCTTACCCAGCATATA 1080  
 Db 1100 GTCCTTTGAAATTAAGGAGAGACACAAAGAAAGAGAGCTTACCCAGCATATA 1159  
 Qy 1081 CTCGCGGAGATTTCTCTCAAGTCAATTTTACCTGGTGTCTGAAATCCAGAGCAATTCCT 1140  
 Db 1160 CTCGCGGAGATTTCTCTCAAGTCAATTTTACCTGGTGTCTGAAATCCAGAGCAATTCCT 1219  
 Qy 1141 AAAAAGCATTTTTCGAGCCCTTGTGAGATTAATCAAGTCAAGTGTGAAAGCGAGG 1200  
 Db 1220 AAAAAGCATTTTTCGAGCCCTTGTGAGATTAATCAAGTCAAGTGTGAAAGCGAGG 1279  
 Qy 1201 CTACAGAGCTGTGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260  
 Db 1280 CTACAGAGCTGTGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1339  
 Qy 1261 TGAGCTGCTGTTGAGATCTCTCTGCTTCCCTTCTTGGCGAGCAACCATAGCTCTC 1320

Db 1340 TGAGCTGCTGTTGAGATCTCTCTGCTTTCCTTCTTTCGAGCAGACACTAGCTCTC 1399  
 Qy 1321 CTCGTCGAACATCTTCTTAACTTCAACCCAGACCATATTTGTTGCAAGCTCAAGTTTA 1380  
 Db 1400 CTCGTCGAACATCTTCTTAACTTCAACCCAGACCATATTTGTTGCAAGCTCAAGTTTA 1459  
 Qy 1381 TTTCAACCCAGAGAGAGCTTCCATTTTGTCTTCAACATTTGTGGAATTTTGTCTACTGCA 1440  
 Db 1460 TTTCAACCCAGAGAGAGCTTCCATTTTGTCTTCAACATTTGTGGAATTTTGTCTACTGCA 1519  
 Qy 1441 ACAGAGTTCTGCGAGAGAGAGATGATGACAGGCTGAGCTTGTGTTGCTTCAAGTT 1500  
 Db 1520 ACAGAGTTCTGCGAGAGAGAGATGATGACAGGCTGAGCTTGTGTTGCTTCAAGTT 1579  
 Qy 1501 CTTCAGCCAAACATACATGATCATCCATGAGAGACAGGAGAGAGAGAGAGAGAGAGAG 1560  
 Db 1580 CTTCAGCCAAACATACATGATCATCCATGAGAGACAGGAGAGAGAGAGAGAGAGAGAG 1639  
 Qy 1561 TCCATCTCTCTCTGAGAACAAATTTCTTCCATTAACAGATGACCCCTCAATCCCATC 1620  
 Db 1640 TCCATCTCTCTCTGAGAACAAATTTCTTCCATTAACAGATGACCCCTCAATCCCATC 1699  
 Qy 1621 ATTAATGTTGGGTCAGAGAACCGGATAGCCCGTTTATTTGGGTTCTTCAACATAGAGAG 1680  
 Db 1700 ATTAATGTTGGGTCAGAGAACCGGATAGCCCGTTTATTTGGGTTCTTCAACATAGAGAG 1759  
 Qy 1681 AAATCTCAAGAACCAACCCAGATGAGAAATTTTGAAGCAATGTGTTTGTGCTGC 1740  
 Db 1760 AAATCTCAAGAACCAACCCAGATGAGAAATTTTGAAGCAATGTGTTTGTGCTGC 1819  
 Qy 1741 AGGATTAAGATGAGGATTTATCTTACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 1800  
 Db 1820 AGGATTAAGATGAGGATTTATCTTACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 1879  
 Qy 1801 ATCTTAATCTAATAAGGTTTCTTCTCAAGAGATGCTCTGTTGGGAGAGAGAGAGAG 1860  
 Db 1880 ATCTTAATCTAATAAGGTTTCTTCTCAAGAGATGCTCTGTTGGGAGAGAGAGAGAG 1939  
 Qy 1861 CCAGAAAGATGATGACAGAACATCTAGCTTCAATGAGGAGAGAGAGAGAGAGAGAG 1920  
 Db 1940 CCAGAAAGATGATGACAGAACATCTAGCTTCAATGAGGAGAGAGAGAGAGAGAGAG 1999  
 Qy 1921 CTCGAGAGAGAGGAGATTTATGTTGTGAGAGATGCAAGAAATTTGCGAAGATGTA 1980  
 Db 2000 CTCGAGAGAGAGGAGATTTATGTTGTGAGAGATGCAAGAAATTTGCGAAGATGTA 2059  
 Qy 1981 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGAGGTTGAGTTGAAATCTAGAGCAATG 2040  
 Db 2060 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGAGGTTGAGTTGAAATCTAGAGCAATG 2119  
 Qy 2041 AAAAACCCTGGCACTTTAAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2097  
 Db 2120 AAAAACCCTGGCACTTTAAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2176

RESULT 2  
 AR144976  
 LOCUS AR144976 3259 bp DNA linear PAT 08-AUG-2001  
 DEFINITION Sequence 23 from patent US 6210950.  
 ACCESSION AR144976  
 VERSION AR144976.1 GI:15106843  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Unknown.  
 Unclassified.  
 REFERENCES  
 1 (bases 1 to 3259)  
 Johnson, W.G. and Stenroos, E.Scott.  
 TITLE Methods for diagnosing, preventing, and treating developmental  
 disorders due to a combination of genetic and environmental factors  
 JOURNAL Patent: US 6210950-A 23 03-APR-2001;  
 FEATURES  
 source  
 1..3259  
 /organism="unknown"

ORIGIN /mol\_type="unassigned DNA"

Query Match 97.6%; Score 2046; DB 6; Length 3259;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2096; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGAGAGAGGTTCTGTATCTATATGCTACACAGCAGGAGCAGGCAAAAGCCATCGAGAA 60
DB ATGAGAGAGGTTCTGTATCTATATGCTACACAGCAGGAGCAGGCAAAAGCCATCGAGAA 139
QY 61 GAAATATGTAGAGCAGCTGTGTACATGTATTTCTGCAGATCTTCACTGTATTAAGAA 120
DB 140 GAAATGTAGAGCAGCTGTGTACATGTATTTCTGCAGATCTTCACTGTATTAAGAA 199
QY 121 TCCGATATGATGACCTTAATAAACCGAAACAGCTCTCTTGTGTGTGTTCTTACAG 180
DB 200 TCCGATATGATGACCTTAATAAACCGAAACAGCTCTCTTGTGTGTGTTCTTACAG 259
QY 181 GGCACCGGAGACCCACCGACACAGCCGCAAGTTGTTAAGGAATACAGAACAAACA 240
DB 260 GGCACCGGAGACCCACCGACACAGCCGCAAGTTGTTAAGGAATACAGAACAAACA 319
QY 241 CTGCCGGTGAATTTCTTGTCTCACTGCGGTATGGGTACTGSGTCTCGGTATTCAGAA 300
DB 320 CTGCCGGTGAATTTCTTGTCTCACTGCGGTATGGGTACTGSGTCTCGGTATTCAGAA 379
QY 301 TACACTTACTTTTGCAATGGGGGGAAGATTAATGATTAACGACTTCAAGACTTGAGCC 360
DB 380 TACACTTACTTTTGCAATGGGGGGAAGATTAATGATTAACGACTTCAAGACTTGAGCC 439
QY 361 CGGCATTTCATGACCTGGACATGACATGATGATGATGATGATGATGATGATGATGATG 420
DB 440 CGGCATTTCATGACCTGGACATGACATGATGATGATGATGATGATGATGATGATGATG 499
QY 421 CGGTGATTTGCTGAGACTGTGCGCAGCCCTCAGAAACATTTTATGTCAGACAGAGCA 480
DB 500 CGGTGATTTGCTGAGACTGTGCGCAGCCCTCAGAAACATTTTATGTCAGACAGAGCA 559
QY 481 GAGGAGATATAGTGGCGCACTCCCGGTGGCATCACTGACATCTTGAAGACAGACTTTG 540
DB 560 GAGGAGATATAGTGGCGCACTCCCGGTGGCATCACTGACATCTTGAAGACAGACTTTG 619
QY 541 AAGTCAGACTGTACATGATTAATCTCAAGTGCAGCTTCTGAGATTTCCATGATTCAGA 600
DB 620 AAGTCAGACTGTACATGATTAATCTCAAGTGCAGCTTCTGAGATTTCCATGATTCAGA 679
QY 601 AGAAGGATTTCTGAGGTTTGAAGCAAAATGCAGTGAACAGCAACCAATCCATGTTGTA 660
DB 680 AGAAGGATTTCTGAGGTTTGAAGCAAAATGCAGTGAACAGCAACCAATCCATGTTGTA 739
QY 661 ATTGAAGACTTTGAGTCTTCACTTACCCGTTGCGTACCCCACTCTCAAGGCTCTCTG 720
DB 740 ATTGAAGACTTTGAGTCTTCACTTACCCGTTGCGTACCCCACTCTCAAGGCTCTCTG 799
QY 721 AATATTCCTGTTTATCCCGCAAAATATTAACAGTACATGTGAAGAGTCTCTTGGCAG 780
DB 800 AATATTCCTGTTTATCCCGCAAAATATTAACAGTACATGTGAAGAGTCTCTTGGCAG 859
QY 781 GAGGAAGCCAAAGTATCTGTGACCTTCAGCAGATCCAGTTTTCAGGTCCAAATTCAA 840
DB 860 GAGGAAGCCAAAGTATCTGTGACCTTCAGCAGATCCAGTTTTCAGGTCCAAATTCAA 919
QY 841 GCAATTCACCTTACTAAGATGATGCCATTAATAACCACTCTGCTGTGATTAATGACAT 900
DB 920 GCAATTCACCTTACTAAGATGATGCCATTAATAACCACTCTGCTGTGATTAATGACAT 979
QY 901 TCAATATACGACTTTTCTATACGCTTGAGATGCTTCAGCGGATTCGCTCCCAAGT 960
DB 980 TCAATATACGACTTTTCTATACGCTTGAGATGCTTCAGCGGATTCGCTCCCAAGT 1039
QY 961 GATTCGAGGTACAAAGCCTTACTCAAAAGCTCAGGCTTGAGATTAAGAGAGCACTGC 1020
```

```
DB 1040 GATTCGAGGTACAAAGCCTTACTCAAAAGCTCAGGCTTGAGATTAAGAGAGCACTGC 1099
QY 1021 GTCCCTTTGAAAATTAAGGACAGACAAAGAAAGAGAGTACTTATCCCGCATATTA 1080
DB 1100 GTCCCTTTGAAAATTAAGGACAGACAAAGAAAGAGAGTACTTATCCCGCATATTA 1159
QY 1081 CCTGGGGAGTGTCTCTCCAGTTCAATTTTACCTGTGTCTTGAATCCGAGCAATTCCT 1140
DB 1160 CCTGGGGAGTGTCTCTCCAGTTCAATTTTACCTGTGTCTTGAATCCGAGCAATTCCT 1219
QY 1141 AAAAAGGCAATTTTGGAGCCCTTGTGACTATACAGTGAACAGTCTGAAAAAGCCAG 1200
DB 1220 AAAAAGGCAATTTTGGAGCCCTTGTGACTATACAGTGAACAGTCTGAAAAAGCCAG 1279
QY 1201 CTACAGAGCTGTGACGTAAACAGGGGAGGCGATTAATAGCCGCTTGAACAGATGCC 1260
DB 1280 CTACAGAGCTGTGACGTAAACAGGGGAGGCGATTAATAGCCGCTTGTGAAGATGCC 1339
QY 1261 TGTGCTGTGTGTGATCTCTCTGCTTCCCTTCTTGCAGGCAACACTCAGTCTC 1320
DB 1340 TGTGCTGTGTGTGATCTCTCTGCTTCCCTTCCCTTGTGCAGGCAACACTCAGTCTC 1399
QY 1321 CTGCTGAAACATCTTCTTAACCTTCAACCCAGACCATATTCGTGCAAGCTCAAGTTA 1380
DB 1400 CTGCTGAAACATCTTCTTAACCTTCAACCCAGACCATATTCGTGCAAGCTCAAGTTA 1459
QY 1381 TTTCACCCAGAAAGCTCAATTTGTCTCAACATTTGGAATTTCTGCTACTGACACA 1440
DB 1460 TTTCACCCAGAAAGCTCAATTTGTCTTCAACATTTGGAATTTCTGCTACTGACACA 1519
QY 1441 ACAGAGTTCCTGCGAAGGAGATATGTACAGGCTGCGCTTGTGTGTTGCTTCAAGTT 1500
DB 1520 ACAGAGTTCCTGCGAAGGAGATATGTACAGGCTGCGCTTGTGTGTTGCTTCAAGTT 1579
QY 1501 CTTACGCAACATACATCATCCATGAAACACGCGGAAAGCCCTGCTCTTAAGTA 1560
DB 1580 CTTACGCAACATACATCATCCATGAAACACGCGGAAAGCCCTGCTCTTAAGTA 1639
QY 1561 TCCATCTCTCTCGAACCAAAATCTTCCACTTACAGATGACCCCTCAATCCCATC 1620
DB 1640 TCCATCTCTCTCGAACCAAAATCTTCCACTTACAGATGACCCCTCAATCCCATC 1699
QY 1621 ATTAATGTTGGTCTCAGAACCCGCAATAGCCGCTTATTTAGGTTCTTACAACTAGAG 1680
DB 1700 ATTAATGTTGGTCTCAGAACCCGCAATAGCCGCTTATTTAGGTTCTTACAACTAGAG 1759
QY 1681 AAACCTCAAGAACCAACCAAGTGAATTTTGAAGCAATGTGTTTGTGGCTGC 1740
DB 1760 AAACCTCAAGAACCAACCAAGTGAATTTTGAAGCAATGTGTTTGTGGCTGC 1819
QY 1741 AGGCAATAGATAGGATTTATCTATTCAGAAAAAGACTCAGACTTTCCTTAAGCATGG 1800
DB 1820 AGGCAATAGATAGGATTTATCTATTCAGAAAAAGACTCAGACTTTCCTTAAGCATGG 1879
QY 1801 ATCTTAATCATCTAAAGGTTTCTCTCAAGAGATGCTCTGTGTGGGAGAGAGAGCC 1860
DB 1880 ATCTTAATCATCTAAAGGTTTCTCTCAAGAGATGCTCTGTGTGGGAGAGAGAGCC 1939
QY 1861 CCAGCAAGATATGTACAGACCAATCAGCTTCAATGSCCAGCAGGTGGCAGAACTCTC 1920
DB 1940 CCAGCAAGATATGTACAGACCAATCAGCTTCAATGSCCAGCAGGTGGCAGAACTCTC 1999
QY 1921 CTCACAGAAACGCGCATATTTATGTGTGTGAGATGCAAGAAATATGCGCAAGATGTA 1980
DB 2000 CTCACAGAAACGCGCATATTTATGTGTGTGAGATGCAAGAAATATGCGCAAGATGTA 2059
QY 1981 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGTGTGAGTTGAAAACTAGAGCAATG 2040
DB 2060 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGTGTGAGTTGAAAACTAGAGCAATG 2119
QY 2041 AAAACCTGCGCACTTTAAAGAAAGAAAAAGCTTCTCAGATATTTGTGCTATA 2097
DB 2120 AAAACCTGCGCACTTTAAAGAAAGAAAAAGCTTCTCAGATATTTGTGCTATA 2176
```

RESULT 3  
LOCUS AX050463 3259 bp DNA linear PAT 12-JAN-2001  
DEFINITION Sequence 23 from Patent WO0071754.  
ACCESSION AX050463  
VERSION AX050463.1 GI:12226668  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS Johnson, W.G. and Stenroos, E.S.  
TITLE Methods for diagnosing, preventing, and treating developmental  
disorders due to a combination of genetic and environmental factors  
JOURNAL Patent: WO 0071754-A 23 30-NOV-2000;  
University of Medicine and Dentistry of New Jersey (US)  
FEATURES  
Source  
1. .3259  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 97.6%; Score 2046; DB 6; Length 3259;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2096; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ATGAGAGGTTTCTGTTACTATATGCTACACAGAGGAGCAAGCCATCGCAGAA 60  
DB 80 ATGAGAGGTTTCTGTTACTATATGCTACACAGAGGAGCAAGCCATCGCAGAA 139  
QY 61 GAATATGTGAGCAAGCTGTGTACATGATTTTCTGAGATCTTCACTGATTAAGTAA 120  
DB 140 GAATATGTGAGCAAGCTGTGTACATGATTTTCTGAGATCTTCACTGATTAAGTAA 139  
QY 121 TCCGATATGATGACCTTAAACCGAAACGAGCTCTTGTGTGTGTTTCTACACG 180  
DB 200 TCCGATATGATGACCTTAAACCGAAACGAGCTCTTGTGTGTGTTTCTACACG 259  
QY 181 GGCACCCGAGACCCAGCAGACAGCCCGCAAGTTTGTAAAGAAATACGAACCAACA 240  
DB 260 GGCACCCGAGACCCAGCAGACAGCCCGCAAGTTTGTAAAGAAATACGAACCAACA 319  
QY 241 CTGCGGTTGATTTCTTGTGCTACCTGCGGATGGGTTTACGGGTCTCGGATTCAGAA 300  
DB 320 CTGCGGTTGATTTCTTGTGCTACCTGCGGATGGGTTTACGGGTCTCGGATTCAGAA 379  
QY 301 TACACCTACTTTTGGCAATGGGGGAGATATTTGATTAAGACTTCAAGAGCTTGAAGC 360  
DB 380 TACACCTACTTTTGGCAATGGGGGAGATATTTGATTAAGACTTCAAGAGCTTGAAGC 439  
QY 361 CGGATTTCTATGACATCTGACATGCAATGATCTGTAGTTTGAACCTTGTGTGAG 420  
DB 440 CGGATTTCTATGACATCTGACATGCAATGATCTGTAGTTTGAACCTTGTGTGAG 499  
QY 421 CCGTGAATGCTGAGACTCTGGCCAGCCCTCAGAAAGATTTTGGTCAAGAGAGAGCAA 480  
DB 500 CCGTGAATGCTGAGACTCTGGCCAGCCCTCAGAAAGATTTTGGTCAAGAGAGAGCAA 559  
QY 481 GAGAGATTAAGTGCGCAGACTCCCGGTGCAATCCTGATCTTGAAGAGAGAGAGCTTGTG 540  
DB 560 GAGAGATTAAGTGCGCAGACTCCCGGTGCAATCCTGATCTTGAAGAGAGAGAGCTTGTG 619  
QY 541 AAGTCAAGCTGCTCAACATTTGAATCTCAAGTCAAGCTTGTAGATTTCATTCAGGA 600  
DB 620 AAGTCAAGCTGCTCAACATTTGAATCTCAAGTCAAGCTTGTAGATTTCATTCAGGA 679  
QY 601 AGAAGGATTTCTGAGTTTGAAGCAAAATGCAATGCAAGCAACCAATCAATGTTGTA 660  
DB 680 AGAAGGATTTCTGAGTTTGAAGCAAAATGCAATGCAAGCAACCAATCAATGTTGTA 739

QY 661 ATTGAAGACTTTGAGTCTCTCACTTACCCGTTGGTACCCCACTCTCACAGGCTCTCTG 720  
DB 740 ATTGAAGACTTTGAGTCTCTCACTTACCCGTTGGTACCCCACTCTCACAGGCTCTCTG 739  
QY 721 AATATTCCTGTTTACCCTCCAGAAATATTTACAGATCATCTGAGAGAGTCTTTGGCCAG 780  
DB 800 AATATTCCTGTTTACCCTCCAGAAATATTTACAGATCATCTGAGAGAGTCTTTGGCCAG 859  
QY 781 GAGAAAGCCAAATATCTGTGACCTTACAGAGATCCAGTTTTCAGTCCCAATTTCAAG 840  
DB 860 GAGAAAGCCAAATATCTGTGACCTTACAGAGATCCAGTTTTCAGTCCCAATTTCAAG 919  
QY 841 GAGTTCAACTTACTACGAATGATGCCATTAACCACTCTGCTGTAGAAATTTGACATT 900  
DB 920 GAGTTCAACTTACTACGAATGATGCCATTAACCACTCTGCTGTAGAAATTTGACATT 979  
QY 901 TCAAAATCAAGCTTTTCTATACGCTTGAAGAGAGCTTCAAGCTGTATTCGCTCAAGT 960  
DB 980 TCAAAATCAAGCTTTTCTATACGCTTGAAGAGAGCTTCAAGCTGTATTCGCTCAAGT 1039  
QY 961 GATTCTGAGGTACAAAGCCTTACTCCAAAGACTGACCTTGAAGATTAAGAGAGACTGC 1020  
DB 1040 GATTCTGAGGTACAAAGCCTTACTCCAAAGACTGACCTTGAAGATTAAGAGAGACTGC 1099  
QY 1021 GTCTTTTGAATAAAGGAGAGACAAAGAAAGAAAGAGACTTACCCAGCATATA 1080  
DB 1100 GTCTTTTGAATAAAGGAGAGACAAAGAAAGAAAGAGACTTACCCAGCATATA 1159  
QY 1081 CTGCGGAGATGTCTCTCCAGTTTCTTAACTGATTTTAACTGATTTTAAATCCAGCAATTCCT 1140  
DB 1160 CTGCGGAGATGTCTCTCCAGTTTCTTAACTGATTTTAACTGATTTTAAATCCAGCAATTCCT 1219  
QY 1141 AAAAAGCATTTTGTGAGAGCCCTTGTGACTATACAGTACAGTGTCTGTAAGAGAGAG 1200  
DB 1220 AAAAAGCATTTTGTGAGAGCCCTTGTGACTATACAGTACAGTGTCTGTAAGAGAGAG 1279  
QY 1201 CTACAGAGCTGTGAGTAAACAAAGGAGAGCCGATTAAGCCGTTTGTACAGATGCC 1260  
DB 1280 CTACAGAGCTGTGAGTAAACAAAGGAGAGCCGATTAAGCCGTTTGTACAGATGCC 1339  
QY 1261 TGTGCTGCTGTGTGATCTCTCTCTGCTTCCCTTCCCTTCCAGGACACACATCACTC 1320  
DB 1340 TGTGCTGCTGTGTGATCTCTCTCTGCTTCCCTTCCCTTCCAGGACACACATCACTC 1399  
QY 1321 CTGCTGAACATCTTCTTAACTTCAACCCAGACCATATTCGTGTGAGACTCAAGTTTA 1380  
DB 1400 CTGCTGAACATCTTCTTAACTTCAACCCAGACCATATTCGTGTGAGACTCAAGTTTA 1459  
QY 1381 TTTCAACCCAGAAAGCTCCATTTTGTCTTCAACATTTGTGAATTTGTCTACTGCCACA 1440  
DB 1460 TTTCAACCCAGAAAGCTCCATTTTGTCTTCAACATTTGTGAATTTGTCTACTGCCACA 1519  
QY 1441 ACAAGGTTCTGCGGAGAGGAGATATGTAACAGCTGTGCTGCTTGTGTGCTTCAAGTT 1500  
DB 1520 ACAAGGTTCTGCGGAGAGGAGATATGTAACAGCTGTGCTGCTTGTGTGCTTCAAGTT 1579  
QY 1501 CTTCAAGCAAAACATACATGATCCCATGAAGACAGCGGAGAAACCTTGCTCTTAAGATA 1560  
DB 1580 CTTCAAGCAAAACATACATGATCCCATGAAGACAGCGGAGAAACCTTGCTCTTAAGATA 1639  
QY 1561 TCCATCTCTCTTGAACAAACAAATCTTTTCACTTCAACAGATGACCCCTCAATCCCCTATC 1620  
DB 1640 TCCATCTCTCTTGAACAAACAAATCTTTTCACTTCAACAGATGACCCCTCAATCCCCTATC 1699  
QY 1621 ATATAGTGGGTCAAGAAACCGGATATAGCCCTTATTTGGTCTTCAACATTAAGAG 1680  
DB 1700 ATATAGTGGGTCAAGAAACCGGATATAGCCCTTATTTGGTCTTCAACATTAAGAG 1759  
QY 1681 AAATCTCAAGAAACAAACCAAGATGGAATTTTGAAGCAATGTGTTTGTGAGCTGC 1740  
DB 1760 AAATCTCAAGAAACAAACCAAGATGGAATTTTGAAGCAATGTGTTTGTGAGCTGC 1819

```

QY 1741 AGCATTAAGATAGGATATATCTATTCAGAAAAAGCTCAGACATTTCTTAAGCATGG 1800
DB 1820 AGCATTAAGATAGGATATATCTATTCAGAAAAAGCTCAGACATTTCTTAAGCATGG 1879
QY 1801 ATCTTAATCATCTTAAGAGTTTCTTCTCAGAGATGCTCTGTTGGGAGAGAGAGCC 1860
DB 1880 ATCTTAATCATCTTAAGAGTTTCTTCTCAGAGATGCTCTGTTGGGAGAGAGAGCC 1939
QY 1861 CCAGCAAGATATATATATATATATATATATATATATATATATATATATATATATAT 1920
DB 1940 CCAGCAAGATATATATATATATATATATATATATATATATATATATATATATATAT 1999
QY 1921 CTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980
DB 2000 CTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2059
QY 1981 CATGATGCTCTTGTGCAATTAATTAAGCAAAAGAGTTGAGTGAATACTAGAGCATG 2040
DB 2060 CATGATGCTCTTGTGCAATTAATTAAGCAAAAGAGTTGAGTGAATACTAGAGCATG 2119
QY 2041 AAAACCTGGCCACTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2097
DB 2120 AAAACCTGGCCACTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2176

```

## RESULT 4

AF025794

LOCUS AF025794 3259 bp mRNA linear PRI 26-MAR-1998

DEFINITION Homo sapiens methionine synthase reductase (MTRR) mRNA, complete cds.

ACCESSION AF025794

VERSION AF025794.1 GI:2981302

KEYWORDS

SOURCE

ORGANISM

Homosapiens (human)

Homosapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 3259)

Lecierc, D., Wilson, A., Dumas, R., Gafuick, C., Song, D., Mackinnon, D.,

Heng, H.H.Q., Romena, J.M., Scherer, S.W., Rosenblatt, D.S. and

Gravel, R.A.

Cloning and mapping of a cDNA for methionine synthase reductase, a

flavoprotein defective in patients with homocystinuria

Proc. Natl. Acad. Sci. U.S.A. 95 (6), 3059-3064 (1998)

JOURNAL

MEDLINE

PUBMED

2 (bases 1 to 3259)

Lecierc, D.

Direct Submission

Submitted (19-SEP-1997) Human Genetics, McGill University -

Montreal Children's Hospital Research Institute, 4060 Ste-Catherine

West, Montreal, Que H3Z 2Z3, Canada

Sequence update by submitter

Location/Qualifiers

1..3259

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/chromosome="5"

/map="5p15.2-p15.3"

1..3259

/gene="MTRR"

80..2176

/note="B12 vitamin, cblE complementation group"

/codon\_start=1

/product="methionine synthase reductase"

## ORIGIN

Query Match 97.6%; Score 2046; DB 9; Length 3259;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2096; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGAGAGGTTTCTGTTACTATATGCTACACAGCAGGAGAGAGAGAGAGAGAGAGAGAGAG 60
DB 80 ATGAGAGGTTTCTGTTACTATATGCTACACAGCAGGAGAGAGAGAGAGAGAGAGAGAGAG 139
QY 61 GAAATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
DB 140 GAAATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 199
QY 121 TCCGATTAAGTATGACCTTAACCAACAGCTCTCTGTTGTTGTTGTTGTTGTTGTTGTTG 180
DB 200 TCCGATTAAGTATGACCTTAACCAACAGCTCTCTGTTGTTGTTGTTGTTGTTGTTGTTG 259
QY 181 GGCACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
DB 260 GGCACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 319
QY 241 CTGCGGTTGATTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 300
DB 320 CTGCGGTTGATTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 379
QY 301 TACACTTACTTTGCAATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
DB 380 TACACTTACTTTGCAATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 439
QY 361 CGGCAATTTCTATGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 440 CGGCAATTTCTATGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 499
QY 421 CCGTGAATGCTGAGACTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 500 CCGTGAATGCTGAGACTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 559
QY 481 GAGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 560 GAGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 619
QY 541 AAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
DB 620 AAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 679
QY 601 AGAAGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB 680 AGAAGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 739
QY 661 ATTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 740 ATTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 799
QY 721 AATATTCCTGTTTACCCCAAGATATTTTACAGGATCATCTGAGAGAGAGAGAGAGAGAGAG 780
DB 800 AATATTCCTGTTTACCCCAAGATATTTTACAGGATCATCTGAGAGAGAGAGAGAGAGAGAG 859

```

QY 781 GAGGAAAGCCAGATCTGTACTTCAGCAGATCCAGTTTTCAGTGCATTTTCAAG 840  
 DB 860 GAGGAAAGCCAGATCTGTACTTCAGCAGATCCAGTTTTCAGTGCATTTTCAAG 919  
 QY 841 GCGATTCATCTCTAGATGATGCGATTAACCACTGCTGGTGAATTTGGACATT 900  
 DB 920 GCGATTCATCTCTAGATGATGCGATTAACCACTGCTGGTGAATTTGGACATT 979  
 QY 901 TCAATATCAGACTTTTCTCTACAGCTGAGATGCTTACGCTGATCTGCTTAAAGT 960  
 DB 980 TCAATATCAGACTTTTCTCTACAGCTGAGATGCTTACGCTGATCTGCTTAAAGT 1039  
 QY 961 GATTCTGAGTACAAAGCTTACTTCAAAAGCTGAGCTTGAAGATTAAGAGCACTGC 1020  
 DB 1040 GATTCTGAGTACAAAGCTTACTTCAAAAGCTGAGCTTGAAGATTAAGAGCACTGC 1099  
 QY 1021 GTCCTTTTGAATTAAGGAGACACAAAGAGAGAGCTTACCCGACATATA 1080  
 DB 1100 GTCCTTTTGAATTAAGGAGACACAAAGAGAGAGCTTACCCGACATATA 1159  
 QY 1081 CCGCGGAGATGTTCTCTCAGATCATTTTACCTGCTTGAATCCGAGCAATTCCT 1140  
 DB 1160 CCGCGGAGATGTTCTCTCAGATCATTTTACCTGCTTGAATCCGAGCAATTCCT 1219  
 QY 1141 AAAAGGCAATTTTGGAGAGCCCTTGTGACTATCCAGTGAAGTCTGAAAAGCCGAG 1200  
 DB 1220 AAAAGGCAATTTTGGAGAGCCCTTGTGACTATCCAGTGAAGTCTGAAAAGCCGAG 1279  
 QY 1201 CTACAGAGCTGTGACATTAACAAAGGAGAGCCGATATAGCCGCTTGTAGAGAGTGC 1260  
 DB 1280 CTACAGAGCTGTGACATTAACAAAGGAGAGCCGATATAGCCGCTTGTAGAGAGTGC 1339  
 QY 1261 TGTGCTGTGTTGATCTCTCTCTGCTTCCCTTCTTCCAGCCAGCACTCAGTCTC 1320  
 DB 1340 TGTGCTGTGTTGATCTCTCTCTGCTTCCCTTCTTCCAGCCAGCACTCAGTCTC 1399  
 QY 1321 CTGCTGCAATCTTCTCTAACTTCAACCCAGACATATTCGTGTGCAAGCTCAAGTTA 1380  
 DB 1400 CTGCTGCAATCTTCTCTAACTTCAACCCAGACATATTCGTGTGCAAGCTCAAGTTA 1459  
 QY 1381 TTTCACCCAGGAAAGCTTCAATTTTGTCTCAACATTTGATGGAATTTCTGTACAGCCACA 1440  
 DB 1460 TTTCACCCAGGAAAGCTTCAATTTTGTCTCAACATTTGATGGAATTTCTGTACAGCCACA 1519  
 QY 1441 ACAGAGTTCTGCGAAGGAGATGATGATACAGGCTGCGCTTGTGTCTTCAAGT 1500  
 DB 1520 ACAGAGTTCTGCGAAGGAGATGATGATACAGGCTGCGCTTGTGTCTTCAAGT 1579  
 QY 1501 CTTGAGCCAAACATATCATGATCCCATGAAAGACGCGGAAAGCCCTGCTCTTAAGATA 1560  
 DB 1580 CTTGAGCCAAACATATCATGATCCCATGAAAGACGCGGAAAGCCCTGCTCTTAAGATA 1639  
 QY 1561 TCCATCTCTCTCTGCAACAAATTTCTTCCACTTACAGATGACCCCTCAATCCCATC 1620  
 DB 1640 TCCATCTCTCTCTGCAACAAATTTCTTCCACTTACAGATGACCCCTCAATCCCATC 1699  
 QY 1621 ATATATGTTGGTCCAGAAACCGGATAGCCCGCTTATTTGAGTTCTTACAAATGAGAG 1680  
 DB 1700 ATATATGTTGGTCCAGAAACCGGATAGCCCGCTTATTTGAGTTCTTACAAATGAGAG 1759  
 QY 1681 AAATCTCAAGAAACACCCAGATGAAATTTTGGAGCAATGTGTTTGGCTGC 1740  
 DB 1760 AAATCTCAAGAAACACCCAGATGAAATTTTGGAGCAATGTGTTTGGCTGC 1819  
 QY 1741 AGGATATAGATATAGGATTTATCTATTCAGAAAGAGCTTCTTAAAGCATGG 1800  
 DB 1820 AGGATATAGATATAGGATTTATCTATTCAGAAAGAGCTTCTTAAAGCATGG 1879  
 QY 1801 ATCTTAATCTATCTAAAGTTTCTTCTCAGAGATGCTCTGTTGGGAGAGAGAGCC 1860  
 DB 1880 ATCTTAATCTATCTAAAGTTTCTTCTCAGAGATGCTCTGTTGGGAGAGAGAGCC 1939  
 QY 1861 CCAGCAAAATATATGATCAAGACATCCAGCTTATGAGCCAGAGTGGGAGAGATCTC 1920

DB 1940 CCAGCAAAATATATGATCAAGACAAATCCAGCTTCAAGTGGCAGAGATCTC 1999  
 QY 1921 CTCAGAGAGAGGCGCATATTTATGTGTGAGATGCAAAAGATATGCGCAAGATGTA 1980  
 DB 2000 CTCAGAGAGAGGCGCATATTTATGTGTGAGATGCAAAAGATATGCGCAAGATGTA 2059  
 QY 1981 CATGATGCCCTTGTGCAATTAATAGCAAAAGGTTTGAAGTTGAAAACTAGAGCAATG 2040  
 DB 2060 CATGATGCCCTTGTGCAATTAATAGCAAAAGGTTTGAAGTTGAAAACTAGAGCAATG 2119  
 QY 2041 AAAACCTTGGCCTTTAAAGAGAAAGAGTACTTACAGATTTTGTGATATA 2097  
 DB 2120 AAAACCTTGGCCTTTAAAGAGAAAGAGTACTTACAGATTTTGTGATATA 2176  
 RESULT 5  
 AF121214  
 LOCUS 3291 bp mRNA linear PRI 12-DEC-1999  
 DEFINITION Homo sapiens methionine synthase reductase (MTRR) mRNA, complete cds.  
 ACCESSION AF121214  
 VERSION AF121214.1 GI:6561338  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 REFERENCE  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 TITLE Leclerc, D., Odievre, M., Wu, Q., Wilson, A., Huizenga, J.J., Rozen, R., Scherer, S.W., and Gravel, R.A.  
 JOURNAL Molecular cloning, expression and physical mapping of the human methionine synthase reductase gene  
 PUBMED Gene 240 (1), 75-88 (1999)  
 MEDLINE 20033550  
 PUBMED 10564814  
 REFERENCE  
 AUTHORS 2 (bases 1 to 3291)  
 Leclerc, D., Odievre, M., H., Wu, Q., Wilson, A., Huizenga, J.J., Johns, T., Shoudridge, E.A., Rosenblatt, D.S., Scherer, S.W., Rozen, R. and Gravel, R.A.  
 TITLE Direct Submision  
 JOURNAL Submitted (18-JAN-1999) Human Genetics, Montreal Children's Hospital, 4060 Ste-Catherine West, Montreal, Quebec H3Z 2Z3, Canada  
 FEATURES  
 source  
 1. .3291  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /chromosome="5"  
 /map="5p15.2-p15.3; between W11755 and D5S1957"  
 1. .3291  
 /gene="MTRR"  
 1. .30  
 /gene="MTRR"  
 /evidence=not\_experimental  
 31. .2208  
 /gene="MTRR"  
 /EC\_number="2.1.1.135"  
 /function="maintenance of methionine synthase in an active state"  
 /note="alternatively spliced; alternative splicing results in extended 5' sequence compared to Genbank Accession Number AF025794; B12 vitamin; cbf1 complementation group"  
 /codon\_start=1  
 /product="methionine synthase reductase"  
 /protein\_id="AA16876.1"  
 /db\_xref="GI:6561339"  
 /translation="MGASVVRAGARLVEALCSFTVTCLEWRRFLLIATVQGGAKA  
 IAEEMQEAVVHGFSAHLICISDSKDYDKETAPLVVVSTGTGDPPTARKFVE  
 IONQTLPEVFAHLRYGLIGDSSEYTYCNGSKLIDKLOELGAFHFTYTGADDCV  
 GLEIVPEWIMKPELPAIRKHFPSRGQBSGLIPVAPSPASLRTDLYKSLIHIESGV  
 ELIRPDSRKQSEVUKQNAVNSNGNVIEDPESLITRSVPLSGASINIPGLPEY  
 LQVHLQESLQGESQVSVTSADPVPVPSKAVQLTTNDRIKTTLLVELDINSIDPSY



QGDARSVICPNDSSEVOSILQRLQLEDRKHCVLKIKADTKKKGATLPOHI PACS  
LQFIETWCLEIRAI PKKAFRLALVDYTSRLOEBSKRSKQAGADYRFPADAC  
LDLILAPSCOPSLILHELPKLOPRYSKASSLTPHGKLHPFNVILSTATT  
EVLKRGVCTGMALIVASLQPMIHSKSDSGKALAPKISIRPTNSPHLDDBSIP  
IIMVGRGRIAPRIGFLQHRBKLOBOHPGNRGMMLPFGCRKDDYLYFKELRHL  
KNGILHLVVSFRDAPVBEERAPATVDNIQLHQOVARILLQENGIHYCGDAKN  
MAKDVIDALVOJISKEVGEKLEAMKTLATLKEKRKYLDIDMS"

transit\_peptide 31..123  
/gene="MTRR"  
/evidence=not\_experimental  
3'UTR 2209..3291  
/gene="MTRR"  
/evidence=not\_experimental  
polyA\_signal 3246..3251  
/gene="MTRR"  
/evidence=not\_experimental  
polyA\_site 3268  
/gene="MTRR"  
/evidence=experimental  
polyA\_site 3276  
/gene="MTRR"  
/evidence=experimental

ORIGIN

Query Match 97.6%; Score 2046; DB 9; Length 3291;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2096; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGAGAGTTTCTGTTACTATATGCTACACAGCAGGAGCAGGCAAAAGCCATGAGAA 60  
DB 112 ATGAGAGAGTTTCTGTTACTATATGCTACACAGCAGGAGCAGGCAAAAGCCATGAGAA 171  
QY 61 GAAATATGAGAGAACTGTGTACATGATTTTTCGAGATCTTCACTATTAATGAA 120  
DB 172 GAAATATGAGAGAACTGTGTACATGATTTTTCGAGATCTTCACTATTAATGAA 231  
QY 121 TCCGATATGATGACCTTAAACCCGAAACAGCTCTCTTGTGTGTGTTTCTACACG 180  
DB 232 TCCGATATGATGACCTTAAACCCGAAACAGCTCTCTTGTGTGTGTTTCTACACG 291  
QY 181 GGCACCGAGAGACCCGACACAGCCCGCAAGTTGTTAAGAAATACAGAACCAACA 240  
DB 292 GGCACCGAGAGACCCGACACAGCCCGCAAGTTGTTAAGAAATACAGAACCAACA 351  
QY 241 CTGCGGTGATTTCTTGTCTCACTGCGGTATGAGTTTCTGCTCTGCTGATTTAGAA 300  
DB 352 CTGCGGTGATTTCTTGTCTCACTGCGGTATGAGTTTCTGCTCTGCTGATTTAGAA 411  
QY 301 TACACTACTTTTGCATATGAGGAGGAGATTAATTGATTAACGACTTCAAGAGCTTGAGCC 360  
DB 412 TACACTACTTTTGCATATGAGGAGGAGATTAATTGATTAACGACTTCAAGAGCTTGAGCC 471  
QY 361 CCGACTTTCTATGACATGACATGACATGATGATGTGTAGGTTTGAACCTTGTGTTAG 420  
DB 472 CCGACTTTCTATGACATGACATGACATGATGATGTGTAGGTTTGAACCTTGTGTTAG 531  
QY 421 CCGTGTATGCTGAGACTTGGCCAGGCCCTCAGAAAGATTTTAGTCAAGCAGAGACA 480  
DB 532 CCGTGTATGCTGAGACTTGGCCAGGCCCTCAGAAAGATTTTAGTCAAGCAGAGACA 591  
QY 481 GAGAGATATAGTGCSCACTCCCGGTGATCACTGACCTTGTAGAGCAGAGCTTGTG 540  
DB 592 GAGAGATATAGTGCSCACTCCCGGTGATCACTGACCTTGTAGAGCAGAGCTTGTG 651  
QY 541 AAGTCAGAGCTGTACACATGATCTCAAGTCGACTTCTGAGATTGATGATTTGAGGA 600  
DB 652 AAGTCAGAGCTGTACACATGATCTCAAGTCGACTTCTGAGATTGATGATTTGAGGA 711  
QY 601 AAAAAAGATTTGAGGTTTGAAGCAAAATGACAGTGAACAGCAACCAATTCATGTTTGA 660  
DB 712 AAAAAAGATTTGAGGTTTGAAGCAAAATGACAGTGAACAGCAACCAATTCATGTTTGA 771  
QY 661 ATTGAAGACTTGAAGTCTCACTTACCCGTTCCGTACCCCACTCTCAAGAGCTCTGTG 720

DB 772 ATTGAAGACTTGAAGTCTCACTTACCCGTTCCGTACCCCACTCTCAAGAGCTCTGTG 831  
QY 721 AATATTCCTGTTTACCCCAAGATTTTACAGTACATCTGACAGAGCTCTTGCCAG 780  
DB 832 AATATTCCTGTTTACCCCAAGATTTTACAGTACATCTGACAGAGCTCTTGCCAG 891  
QY 781 GAGAAAGCCAGATATCTGTGACTTCAAGATTCAGAGTTCAGTTCAGTTCAGTTCAG 840  
DB 892 GAGAAAGCCAGATATCTGTGACTTCAAGATTCAGAGTTCAGTTCAGTTCAGTTCAG 951  
QY 841 GAGATTCAGCTTACAGAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
DB 952 GAGATTCAGCTTACAGAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1011  
QY 901 TCAATATCAGACTTTTCTCTATCAGCTGAGATGCTTCAAGGTGATCTGCTTCAAGT 960  
DB 1012 TCAATATCAGACTTTTCTCTATCAGCTGAGATGCTTCAAGGTGATCTGCTTCAAGT 1071  
QY 961 GATTCGAGTACAAAGCTTACCAAGACTGACAGCTTGAAGATTAAGAGAGCACTGC 1020  
DB 1072 GATTCGAGTACAAAGCTTACCAAGACTGACAGCTTGAAGATTAAGAGAGCACTGC 1131  
QY 1021 GTCTTTTGAATAAAGCAGACACAAAGAAAGAGCTTACCTTACCCAGCATATA 1080  
DB 1132 GTCTTTTGAATAAAGCAGACACAAAGAAAGAGCTTACCTTACCCAGCATATA 1191  
QY 1081 CCGCGGAGATGTTCTCTCAAGTTCATTTTACCTGTGTCTTGAATCCGAGCAATTCCT 1140  
DB 1192 CCGCGGAGATGTTCTCTCAAGTTCATTTTACCTGTGTCTTGAATCCGAGCAATTCCT 1251  
QY 1141 AAAAAAGCAATTTTGGAGCCCTGTGACATPACAGAGACAGTCTGAAAAGGCGAG 1200  
DB 1252 AAAAAAGCAATTTTGGAGCCCTGTGACATPACAGAGACAGTCTGAAAAGGCGAG 1311  
QY 1201 CTACAGAGCTGTGACATTAACAAGGAGGAGCGATTAAGCGCTTGTACAGATGCG 1260  
DB 1312 CTACAGAGCTGTGACATTAACAAGGAGGAGCGATTAAGCGCTTGTACAGATGCG 1371  
QY 1261 TGTGCTGCTGTTGATTCCTCTCTGCTGCTTCTCTGCTGCTGCTGCTGCTGCTGCT 1320  
DB 1372 TGTGCTGCTGTTGATTCCTCTCTGCTGCTTCTCTGCTGCTGCTGCTGCTGCTGCT 1431  
QY 1321 CTGCTCGAATCTTCTCTTAACTTCAACCAAGCAATGAGGAGGAGGAGGAGGAGGAG 1380  
DB 1432 CTGCTCGAATCTTCTCTTAACTTCAACCAAGCAATGAGGAGGAGGAGGAGGAGGAG 1491  
QY 1381 TTTCAACCAAGAAAGCTCCATTTTGTCTTCAATTTGTGAAATTTCTGTCTAGCCACA 1440  
DB 1492 TTTCAACCAAGAAAGCTCCATTTTGTCTTCAATTTGTGAAATTTCTGTCTAGCCACA 1551  
QY 1441 ACGAGGTTCTCGAGAGGAGATATGACAGGCTGAGCTGCTGTTGTTGTTGCTTCAAGTT 1500  
DB 1552 ACGAGGTTCTCGAGAGGAGATATGACAGGCTGAGCTGCTGTTGTTGTTGCTTCAAGTT 1611  
QY 1501 CTTGAGCCAAATATCAGTCCATGCCATGAAAGCAGCGGAAAGCCCTGCTCTTAAAGTA 1560  
DB 1612 CTTGAGCCAAATATCAGTCCATGCCATGAAAGCAGCGGAAAGCCCTGCTCTTAAAGTA 1671  
QY 1561 TCCATCTCTCTGAAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
DB 1672 TCCATCTCTCTGAAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1731  
QY 1621 ATTAATGTTGGTTCAGAAACCGGATAGCCCGGTTTATTTGAGTTCTTACAAATAGAG 1680  
DB 1732 ATTAATGTTGGTTCAGAAACCGGATAGCCCGGTTTATTTGAGTTCTTACAAATAGAG 1791  
QY 1681 AATCTCCAAAGCAACACCCAGATGAAATTTTGAAGCAATGAGTGTGTTTGTGCTGC 1740  
DB 1792 AATCTCCAAAGCAACACCCAGATGAAATTTTGAAGCAATGAGTGTGTTTGTGCTGC 1851  
QY 1741 AGCATATGAGATGAGGATTTATCTATTCAGAAAGAGCTCAGACATTTCTTAAAGATGG 1800



Db	1852	AGCATAAGATAGGAGTTATCTATTTCAGAAAAAGCTCAACACTTTCCTTAAGATGG	1911
Qy	1801	ATCTTAATCTACTTAAAGGTTTCCTCTTCAAGAGATGCTCTGTTGGGAGAGAAACC	1860
Db	1912	ATCTTAATCTACTTAAAGGTTTCCTCTTCAAGAGATGCTCTGTTGGGAGAGAAACC	1971
Qy	1861	CCAGCAAGTATGTATCAAGACAACTCCAGCTTCATGGCCAGCAGGTGGCGAATCCTC	1920
Db	1972	CCAGCAAGTATGTATCAAGACAACTCCAGCTTCATGGCCAGCAGGTGGCGAATCCTC	2031
Qy	1921	CTCCAGAGAAACGGCCATATTTATGTGTGGAGATGCAAGAAATATATGCGCAAGATGTA	1980
Db	2032	CTCCAGAGAAACGGCCATATTTATGTGTGGAGATGCAAGAAATATATGCGCAAGATGTA	2091
Qy	1981	CATGATGCCCTTGTGCAATATATATAGCAAGAGGTTGGAGTTGAAAACTAGAAACATG	2040
Db	2092	CATGATGCCCTTGTGCAATATATATAGCAAGAGGTTGGAGTTGAAAACTAGAAACATG	2151
Qy	2041	AAAACCTGGCCACTTTTAAAAAGAGAAAAACGCTACCTTCAGAGATATTTGGTCATAA	2097
Db	2152	AAAACCTGGCCACTTTTAAAAAGAGAAAAACGCTACCTTCAGAGATATTTGGTCATAA	2208
RESULT 6			
LOCUS	BC054816	3310 bp	mRNA
DEFINITION			linear
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
1 (bases 1 to 3310)			
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			
31			
32			
33			
34			
35			
36			
37			
38			
39			
40			
41			
42			
43			
44			
45			
46			
47			
48			
49			
50			
51			
52			
53			
54			
55			
56			
57			
58			
59			
60			
61			
62			
63			
64			
65			
66			
67			
68			
69			
70			
71			
72			
73			
74			
75			
76			
77			
78			
79			
80			
81			
82			
83			
84			
85			
86			
87			
88			
89			
90			
91			
92			
93			
94			
95			
96			
97			
98			
99			
100			
101			
102			
103			
104			
105			
106			
107			
108			
109			
110			
111			
112			
113			
114			
115			
116			
117			
118			
119			
120			
121			
122			
123			
124			
125			
126			
127			
128			
129			
130			
131			
132			
133			
134			
135			
136			
137			
138			
139			
140			
141			
142			
143			
144			
145			
146			
147			
148			
149			
150			
151			
152			
153			
154			
155			
156			
157			
158			
159			
160			
161			
162			
163			
164			
165			
166			
167			
168			
169			
170			
171			
172			
173			
174			
175			
176			
177			
178			
179			
180			
181			
182			
183			
184			
185			
186			
187			
188			
189			
190			
191			
192			
193			
194			
195			
196			
197			
198			
199			
200			
201			
202			
203			
204			
205			
206			
207			
208			
209			
210			
211			
212			
213			
214			
215			
216			
217			
218			
219			
220			
221			
222			
223			
224			
225			
226			
227			
228			
229			
230			
231			
232			
233			
234			
235			
236			
237			
238			
239			
240			
241			
242			
243			
244			
245			
246			
247			
248			
249			
250			
251			
252			
253			
254			
255			
256			
257			
258			
259			
260			
261			
262			
263			
264			
265			
266			
267			
268			
269			
270			
271			
272			
273			
274			
275			
276			
277			
278			
279			
280			
281			
282			
283			
284			
285			
286			
287			
288			
289			
290			
291			
292			
293			
294			
295			
296			
297			
298			
299			
300			
301			
302			
303			
304			
305			
306			
307			
308			
309			
310			
311			
312			
313			
314			
315			
316			
317			
318			
319			
320			
321			
322			
323			
324			
325			
326			
327			
328			
329			
330			
331			
332			
333			
334			
335			
336			
337			
338			
339			
340			
341			
342			
343			
344			
345			
346			
347			
348			
349			
350			
351			
352			
353			
354			
355			
356			
357			
358			
359			
360			
361			
362			
363			
364			
365			
366			
367			
368			
369			
370			
371			
372			
373			
374			
375			
376			
377			
378			
379			
380			
381			
382			
383			
384			
385			
386			
387			
388			
389			
390			
391			
392			
393			
394			
395			
396			
397			
398			
399			
400			
401			
402			
403			
404			
405			
406			
407			
408			
409			
410			
411			
412			
413			
414			
415			
416			
417			
418			
419			
420			
421			
422			
423			
424			
425			
426			
427			
428			
429			
430			
431			
432			
433			
434			
435			
436			
437			
438			
439			
440			
441			
442			
443			
444			
445			
446			
447			
448			
449			
450			
451			
452			
453			
454			
455			
456			
457			
458			
459			
460			
461			
462			
463			
464			
465			
466			
467			
468			
469			
470			
471			
472			
473			
474			
475			
476			
477			
478			
479			
480			
481			</

DNA sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: <http://www.nisc.nih.gov/> Contact: nisc.mc@nih.gov

Ahter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakeley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karling, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McLooney, J.C., McDowell, J., Pearson, R., Statistop, S., Thomas, P.J., Touchman, J.W., Taugenon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLAT at: <http://image.llnl.gov> Series: IRAC Plate: 115 Row: d Column: 11 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4505278.

Location/Qualifiers

1. .3310

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5205285"

/issue\_type="Lung, Splen, fetal, pooled"

/clone\_lib="NIH MGC\_122"

/lab\_host="DH10B"

/note="Vector: pCMV-SPORT6"

<1. .3310

/gene="MTRR"

/note="synonym: MSR"

/db\_xref="LOCUSID:4552"

/db\_xref="MIM:602568"

<1. .2149

/gene="MTRR"

/codon\_start=2

/product="MTRR protein"

/protein\_id="AAH54816.1"

/db\_xref="GI:33392776"

/db\_xref="LOCUSID:4552"

/db\_xref="MIM:602568"

/translation="VRRLARLRGFTVTCLEVMRFLIVATGOGAKAIEBICEAVVHGSDADHICISDSKDYDLKTEPAFLVVVSTGTGPPPTARKFVEIONQLPDPFPAHLRYGLIGLGESEYVPCNGGKILDKRLOEGARHFTGHAADCVGLGLVVEPIAGLWPLAKRHRSSRGQEBISGALPVASPASSTDLVKSLLHIESQVLELRPDSGRKDSERVAKONAVNOSQSVIVYEDRESSITRESVPLSQSLNIPGLPEYLOVHIOESLQESQSVTSADPVFQVPLISKAVOYLTNDIKITTLVELDINSITDESIOGDASVYCPSNDSEVQSLQRLQEDKREKCVLLIKADTKKKGATLPOHLPAGSLQIFTCWLEIRAI PKKAFPLRPLVDYSDSAEKRRLOELSGKQGANYSFVDAACLDLILAFPSGCPPLSLLEHLPLTLOPRPYSCASSLFHGKGLFVNFIVFELSTATTEVLARGVGTGAPLILVAVSLOPNTHASHEDSGKLAIPKISLPTTNSLIVELHIESQVLELRPDSGRAPFSGFQHKREKIQEQRPDNFGAMTLFPCGRKDPVLRKELRHLKIGLITHLKVSPFSDAVGEEBAFPAKTVQDNTQDHGGQVARIILIOENGHIIYVGDAKNAKVDHALVQIISKEVGEKLEAMKTLATLKEKRYLDIDWS"

59. .493

/gene="MTRR"

/note="flavodoxin; Region: Flavodoxin"

/db\_xref="CDD:pfam00258"

848. .2143

/gene="MTRR"

/note="CysJ; Region: Sulfite reductase, alpha subunit (flavoprotein) [inorganic ion transport and metabolism]"

/db\_xref="CDD:COG0369"

ORIGIN

Query Match 95.1%; Score 1995; DB 9; Length 3310;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 2095; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGAGAGAGTTCTGTACTATATAGTCACACAGCAGGACGACGAAAGCCATGCGACAA 60

53 ATGAGAGAGTTCTGTACTATATAGTCACACAGCAGGACGACGAAAGCCATGCGACAA 112

61 GAAATATGAGCAAGCTGTGGTACATGATTTTCTGCAGATCTTCACTGATTAAGTGA 120  
113 GAAATATGAGCAAGCTGTGGTACATGATTTTCTGCAGATCTTCACTGATTAAGTGA 172  
121 TCCGATTAAGTATGACCTTAATAAAGCAAGCTCTCTTGTGTGTGTGTCTTCAACAG 180  
173 TCCGATTAAGTATGACCTTAATAAAGCAAGCTCTCTTGTGTGTGTGTCTTCAACAG 232  
181 GGCACCGGAGACCCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACA 240  
233 GGCACCGGAGACCCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACA 292  
241 CTGCGGATGATTTCTTGTCTCACTGCGGATGATGATGATGATGATGATGATGAT 300  
293 CTGCGGATGATTTCTTGTCTCACTGCGGATGATGATGATGATGATGATGATGAT 352  
301 TACACCTACTTTTGCATGAGGAGGAGATTAATTAATTAATTAATTAATTAATTAAT 360  
353 TACACCTACTTTTGCATGAGGAGGAGATTAATTAATTAATTAATTAATTAATTAAT 412  
361 CGGCAATTTCTATGACATGACATGACATGACATGACATGACATGACATGACATGAC 420  
413 CGGCAATTTCTATGACATGACATGACATGACATGACATGACATGACATGACATGAC 472  
421 CGGCAATTTCTATGACATGACATGACATGACATGACATGACATGACATGACATGAC 480  
473 CGGCAATTTCTATGACATGACATGACATGACATGACATGACATGACATGACATGAC 532  
481 GAGAGATTAAGTGTGCGCATCTCCGAGTGTGATCATCTGATCTTGAAGACAGACCTTGTG 540  
533 GAGAGATTAAGTGTGCGCATCTCCGAGTGTGATCATCTGATCTTGAAGACAGACCTTGTG 592  
541 AAGTCAGAGCTGTGACATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600  
593 AAGTCAGAGCTGTGACATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 652  
601 AAGTCAGAGCTGTGACATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 660  
653 AAGTCAGAGCTGTGACATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 712  
661 AAGTCAGAGCTGTGACATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 720  
713 AAGTCAGAGCTGTGACATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 772  
721 AAGTCAGAGCTGTGACATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 780  
773 AAGTCAGAGCTGTGACATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 832  
781 GAGAGAAAGCAAGTATCTGTGACATGACATGACATGACATGACATGACATGACATGAC 840  
833 GAGAGAAAGCAAGTATCTGTGACATGACATGACATGACATGACATGACATGACATGAC 892  
841 GAGAGAAAGCAAGTATCTGTGACATGACATGACATGACATGACATGACATGACATGAC 900  
893 GAGAGAAAGCAAGTATCTGTGACATGACATGACATGACATGACATGACATGACATGAC 952  
901 TCAAAATACAGACTTTTCTATCAGCTGTGAGATGCTTCAAGCTGTGAGATGCTTCAAG 960  
953 TCAAAATACAGACTTTTCTATCAGCTGTGAGATGCTTCAAGCTGTGAGATGCTTCAAG 1012  
961 GATTCGAGGTACAAAGCTTACCTCAAGATGACATGACATGACATGACATGACATGACAT 1020  
1013 GATTCGAGGTACAAAGCTTACCTCAAGATGACATGACATGACATGACATGACATGACAT 1072  
1021 GTCCTTTTGAATAAAGGACACAAAGAAAGAAAGAGCTTACCTTACCTTACCTTACCT 1080  
1073 GTCCTTTTGAATAAAGGACACAAAGAAAGAAAGAGCTTACCTTACCTTACCTTACCT 1132  
1081 CCGCGGAGATGTTCTCTCAGATCAATTTTACCTGCTGTGAAATTCGAGCAATTTCT 1140  
1133 CCGCGGAGATGTTCTCTCAGATCAATTTTACCTGCTGTGAAATTCGAGCAATTTCT 1192

1141 AAAAAAGCATTTTTCGAGCCCTTGTGACATTAACAGTACAGTGTGAAAAAGCGCAGG 1200  
1193 AAAAAAGCATTTTTCGAGCCCTTGTGACATTAACAGTACAGTGTGAAAAAGCGCAGG 1252  
1201 CTACAGAGCTGTGACATTAACAGAGGAGAGGAGTATTAAGCCGCTTTTGAAGATGAC 1260  
1253 CTACAGAGCTGTGACATTAACAGAGGAGAGGAGTATTAAGCCGCTTTTGAAGATGAC 1312  
1261 TGTGCTGCTGTGTGATCTCTCTGCTGCTTCCCTTCTGACAGCAGCAGTCAAGTCTC 1320  
1313 TGTGCTGCTGTGTGATCTCTCTGCTGCTTCCCTTCTGACAGCAGCAGTCAAGTCTC 1372  
1321 CTGCTGGAACATCTTCTAACTTCAACCCAGACCATATGCTGTGCAAGCTCAAGTTTA 1380  
1373 CTGCTGGAACATCTTCTAACTTCAACCCAGACCATATGCTGTGCAAGCTCAAGTTTA 1432  
1381 TTTCAACCCAGAGAAAGCTCATTCTTCAACATTTGCAATTTTCTGCTACTGACACA 1440  
1433 TTTCAACCCAGAGAAAGCTCATTCTTCAACATTTGCAATTTTCTGCTACTGACACA 1492  
1441 ACAGAGGTTCTGCGAGAGGAGATGTAACAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1500  
1493 ACAGAGGTTCTGCGAGAGGAGATGTAACAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1552  
1501 CTTCAGCCAAACATACATGATCCATGAGACAGAGGAGAAAGCCCTGAGCTCTAAGATA 1560  
1553 CTTCAGCCAAACATACATGATCCATGAGACAGAGGAGAAAGCCCTGAGCTCTAAGATA 1612  
1561 TCCATCTCTCTGAGAACAAATTTCTTCACTTACAGATGACCCCTGATCCCATC 1620  
1613 TCCATCTCTCTGAGAACAAATTTCTTCACTTACAGATGACCCCTGATCCCATC 1672  
1621 AATATGAGGTTCTGCGAGAGGAGATGTAACAGCTGAGCTGAGCTGAGCTGAGCTGAG 1680  
1673 AATATGAGGTTCTGCGAGAGGAGATGTAACAGCTGAGCTGAGCTGAGCTGAGCTGAG 1732  
1681 AATATGAGGTTCTGCGAGAGGAGATGTAACAGCTGAGCTGAGCTGAGCTGAGCTGAG 1740  
1733 AATATGAGGTTCTGCGAGAGGAGATGTAACAGCTGAGCTGAGCTGAGCTGAGCTGAG 1792  
1741 AAGCATTAAGGATTAAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1800  
1793 AAGCATTAAGGATTAAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1852  
1801 AATCTTAACATCTTAAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1860  
1853 AATCTTAACATCTTAAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1912  
1861 CCAGCAAGTATGTAACAAGACATCCAGCTTCAATGAGCAGAGTGTGAGAAATCTCTC 1920  
1913 CCAGCAAGTATGTAACAAGACATCCAGCTTCAATGAGCAGAGTGTGAGAAATCTCTC 1972  
1921 CTTCAGAGAAAGGCGCATTTTATGTGTGTGAGATGCAAGAAATATGAGCAAGATGTA 1980  
1973 CTTCAGAGAAAGGCGCATTTTATGTGTGTGAGATGCAAGAAATATGAGCAAGATGTA 2032  
1981 CATGATGCTCTGTGTGCAATTAATTAAGCAAGAGGTTGTGAAATACTTAAGAAAGTA 2040  
2033 CATGATGCTCTGTGTGCAATTAATTAAGCAAGAGGTTGTGAAATACTTAAGAAAGTA 2092  
2041 AAAACCTGAGCACTTAATAAGAAAGAAAGCTTACAGATTAATTTGTGATATA 2097  
2093 AAAACCTGAGCACTTAATAAGAAAGAAAGCTTACAGATTAATTTGTGATATA 2149

RESULT 7  
BV177620 2933 bp DNA linear STS 10-JUN-2004  
LOCUS BV177620  
DEFINITION egm95800 Human DNA (Sequenc) Homo sapiens STS genomic, sequence  
tagged site.  
ACCESSION BV177620  
VERSION BV177620.1 GI:48013757  
KEYWORDS STS.

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 2933)  
AUTHORS Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,  
TITLE Large-Scale Validation of Single Nucleotide Polymorphisms in Gene  
JOURNAL Regions  
COMMENT Genome Res. (2004) In press  
Contact: Andreas Braun  
Pharmaceuticals division  
Sequenom, Inc.  
3595 John Hopkins Court, San Diego, CA 92121, USA  
Tel: 18582029018  
Fax: 18582029020  
Email: abraun@sequenom.com  
Primer A: No primer sequence submitted  
Primer B: No primer sequence submitted  
STS size: 2933.  
FEATURES  
source Location/Qualifiers  
1..2933  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone\_lib="Human DNA (Sequenom)"  
<1..2933  
ORIGIN  
Query Match 63.8%; Score 1338; DB 11; Length 2933;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1778; Conservative 0; Mismatches 4; Indels 2; Gaps 2;  
QY 315 CAATGGGGGAAGATTAATGTAATAACGACTTCAAGACTTGAAGCCCGGCAATTTCTATGA 374  
Db 2846 CAATGGGGGAAGATTAATGTAATAACGACTTCAAGACTTGAAGCCCGGCAATTTCTATGA 2787  
QY 375 CACTGCATGCAGATGACTGTGTAGGTTTGAACCTTGTGGTGAAGCCGTGATGCTGG 434  
Db 2786 CACTGCATGCAGATGACTGTGTAGGTTTGAACCTTGTGGTGAAGCCGTGATGCTGG 2727  
QY 435 ACTCTGGCCAGCCCTCGAAGAGCATTTTAGTCAAGAGAGACAAGAGAGATTAAGTGG 494  
Db 2726 ACTCTGGCCAGCCCTCGAAGAGCATTTTAGTCAAGAGAGACAAGAGAGATTAAGTGG 2667  
QY 495 CGCACTCCCGGTGGCATCACTGCATCTTGAAGACAGACCTTGTGAAGTCAAGACTGCT 554  
Db 2666 CGCACTCCCGGTGGCATCACTGCATCTTGAAGACAGACCTTGTGAAGTCAAGACTGCT 2607  
QY 555 ACAATTGAATCTCAAGTGAAGCTTCTGAATTCGATGATTCAGAAAGAAAGATTCTGA 614  
Db 2606 ACAATTGAATCTCAAGTGAAGCTTCTGAATTCGATGATTCAGAAAGAAAGATTCTGA 2547  
QY 615 GCTTTTGAAGCAAAATGCAAGTGAACAGAACCAATCCAATGTTGTAATTGAAGCTTGA 674  
Db 2546 GCTTTTGAAGCAAAATGCAAGTGAACAGAACCAATCCAATGTTGTAATTGAAGCTTGA 2487  
QY 675 GTCCTCACTTACCGCTCGGTACCCCACTCTCAAGACCTCTCTGAATATTCCTGGTTT 734  
Db 2486 GTCCTCACTTACCGCTCGGTACCCCACTCTCAAGACCTCTCTGAATATTCCTGGTTT 2427  
QY 735 ACCCCAGAAATATTACAGGTACATCTGCAGAGACTCTTGGCCAGAGAGAAAGCCAAAGT 794  
Db 2426 ACCCCAGAAATATTACAGGTACATCTGCAGAGACTCTTGGCCAGAGAGAAAGCCAAAGT 2367  
QY 795 ATCTGTGATCTTCAGCAATCCAGTTTTCAGTGCCTAATTCAGAGCAATTCACCTTAC 854  
Db 2366 ATCTGTGATCTTCAGCAATCCAGTTTTCAGTGCCTAATTCAGAGCAATTCACCTTAC 2307  
QY 855 TAGCAATGATGCCATAAAGAACCACTGTGTGTGAATGGAATTGACATTTCAATACAGATT 914  
Db 2306 TAGCAATGATGCCATAAAGAACCACTGTGTGTGAATGGAATTGACATTTCAATACAGATT 2247

QY 915 TTCCTATCAGCCTGGAGATGCTTCAAGCGTGAATCTGCCCTTAACAGTATTTCTGAGGTACA 974  
Db 2246 TTCCTATCAGCCTGGAGATGCTTCAAGCGTGAATCTGCCCTTAACAGTATTTCTGAGGTACA 2187  
QY 975 AAGCCCTACTCCAAAGCTGACGCTTGAAGATTAAGAGAGCACTGGCTCTTTGAAAT 1034  
Db 2186 AAGCCCTACTCCAAAGCTGACGCTTGAAGATTAAGAGAGCACTGGCTCTTTGAAAT 2127  
QY 1035 AAAGGACAGACAAAGAGAAAGAGAGCTTACCCAGCATATACCTGGGGATGTTT 1094  
Db 2126 AAAGGACAGACAAAGAGAAAGAGAGCTTACCCAGCATATACCTGGGGATGTTT 2067  
QY 1095 TCTCCAGTTCAATTTTAACTGCTGTGTGTAATCCAGCAATTCCTTAAAGGCAATTTT 1154  
Db 2066 TCTCCAGTTCAATTTTAACTGCTGTGTGTAATCCAGCAATTCCTTAAAGGCAATTTT 2007  
QY 1155 GCGAGCCCTTGGAGCATATACAGTACAGTGCAGTGTGAAAAGCCGAGGCTACAGAGCTGTG 1214  
Db 2006 GCGAGCCCTTGGAGCATATACAGTACAGTGCAGTGTGAAAAGCCGAGGCTACAGAGCTGTG 1947  
QY 1215 CAGTAAACAAGGGGACGCCGATTAATAGCCGCTTGTAGCAGAGATGCTGTGCTGTGTT 1274  
Db 1946 CAGTAAACAAGGGGACGCCGATTAATAGCCGCTTGTAGCAGAGATGCTGTGCTGTGTT 1887  
QY 1275 GGAATCTCTCTGCTTCTTCCCTTCTTGGCAAGCCACCACTGATCTCTGCTGAACATCT 1334  
Db 1886 GGAATCTCTCTGCTTCTTCCCTTCTTGGCAAGCCACCACTGATCTCTGCTGAACATCT 1827  
QY 1335 TCCTAAACCTTCAACCCAGACCAATATGATGTGTAAGCTCAAGTTTATTCACCAGAGAAA 1394  
Db 1826 TCCTAAACCTTCAACCCAGACCAATATGATGTGTAAGCTCAAGTTTATTCACCAGAGAAA 1767  
QY 1395 GCTCCATTTTGTCTTCAACATGTGGAATTTCTGTACTGCAACAAGAGGTTTGGC 1454  
Db 1766 GCTCCATTTTGTCTTCAACATGTGGAATTTCTGTACTGCAACAAGAGGTTTGGC 1707  
QY 1455 GAAAGGATGTGATCAGGCTGCTGCTGCTTGTGTTGCTTCAAGTTCTTCAAGCCAAACAT 1514  
Db 1706 GAAAGGATGTGATCAGGCTGCTGCTGCTTGTGTTGCTTCAAGTTCTTCAAGCCAAACAT 1647  
QY 1515 ACATGATCCCAATGAAGACAGC--GGGAAAAGCCCTGCTCCTTAATATTCATCTCTCCTC 1573  
Db 1646 ACATGATCCCAATGAAGACAGC--GGGAAAAGCCCTGCTCCTTAATATTCATCTCTCCTC 1587  
QY 1574 GAACAACAATTTCTTCACTTACAGATGACCCCTCAATCCCATCATATATGATGCTGTC 1633  
Db 1586 GAACAACAATTTCTTCACTTACAGATGACCCCTCAATCCCATCATATATGATGCTGTC 1527  
QY 1634 CAGGAACCGGATAGCCCGGTTTATGGGTTCTTACAACTTAGAGAAACTCCAGAAC 1693  
Db 1526 CAGGAACCGGATAG--CCCGTTTATGGGTTCTTACAACTTAGAGAAACTCCAGAAC 1468  
QY 1694 AACACCCAGATGGAATTTTGGAGCAATGATGTTTGGCTGCTGAGGATTAAGATA 1753  
Db 1467 AACACCCAGATGGAATTTTGGAGCAATGATGTTTGGCTGCTGAGGATTAAGATA 1408  
QY 1754 GGGATTATCTATCAGAAAAGAGCTCAGACATTTCTTAAAGATGGATTTTAATCATC 1813  
Db 1407 GGGATTATCTATCAGAAAAGAGCTCAGACATTTCTTAAAGATGGATTTTAATCATC 1348  
QY 1814 TAAAGTTTCTTCTCAAGATGCTCTGTTGGGAGAGAGAAAGCCCGCAAGATATG 1873  
Db 1347 TAAAGTTTCTTCTCAAGATGCTCTGTTGGGAGAGAGAAAGCCCGCAAGATATG 1288  
QY 1874 TACAAGCAACATCAGCTTCAATGGCAGAGGATGAGAAATCTCTCCAGAGAAAG 1933  
Db 1287 TACAAGCAACATCAGCTTCAATGGCAGAGGATGAGAAATCTCTCCAGAGAAAG 1228  
QY 1934 GCCATATTTATGTGTGTGAGATGCAAGAAATATGCGCAAGATGATCATGATGCTCTTG 1993  
Db 1227 GCCATATTTATGTGTGTGAGATGCAAGAAATATGCGCAAGATGATCATGATGCTCTTG 1168

QY 1994 TGCATAATAAAGCAAGAGTTGAGTTGAAAACTAGAGCAATGAAAAACCTTGCCA 2053  
DB 1167 TCGAATAATATTAAGCAAGAGTTGAGTTGAAAACTAGAGCAATGAAAAACCTTGCCA 1108  
QY 2054 CTTTAAAGAGAAAAACCTTACCTTCAGATTTTGTGATTA 2097  
DB 1107 CTTTAAAGAGAAAAACCTTACCTTCAGATTTTGTGATTA 1064

RESULT 8  
BV178010/c 2933 bp DNA linear STS 10-JUN-2004  
LOCUS sqm97986 Human DNA (Sequenom) Homo sapiens STS genomic, sequence  
DEFINITION tagged site.  
ACCESSION BV178010  
VERSION BV178010.1 GI:48014252  
KEYWORDS STS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 2933)  
Nelson,R.M., Marnell, G., Kammerer, S., Hoyal, C.R., Shi, M.M.,  
Cantor, C.R. and Braun, A.  
Large-Scale Validation of Single Nucleotide Polymorphisms in Gene  
Regions  
Genome Res. (2004) In press

JOURNAL COMMENT  
Contact: Andreas Braun  
Pharmaceuticals division  
Sequenom, Inc.  
3595 John Hopkins Court, San Diego, CA 92121, USA  
Tel: 18582029018  
Fax: 18582029020  
Email: abraun@sequenom.com  
Primer A: No primer sequence submitted  
Primer B: No primer sequence submitted  
STS size: 2933.

FEATURES  
source Location/Qualifiers  
1..2933  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone\_lib="Human DNA (Sequenom)"  
1..>2933

ORIGIN  
STS  
Query Match 63.8%; Score 1338; DB 11; Length 2933;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1778; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

DB 2546 GGTTTTGAAGCAAAATGCAAGTGAACAGCAACCAATTCAAATGTTGTAATTGAACCTTTGA 2487  
QY 675 GTCTCACTTACCCGTTGGGTACCCCACTCTCAAGAGCTCTGTAATATTCGTGTT 734  
DB 2486 GTCTCACTTACCCGTTGGGTACCCCACTCTCAAGAGCTCTGTAATATTCGTGTT 2427  
QY 735 ACCCCAGAAATTTTACAGATTCATCTGCAAGAGTCTCTTGGCCAGAGAAAGCCAGT 794  
DB 2426 ACCCCAGAAATTTTACAGATTCATCTGCAAGAGTCTCTTGGCCAGAGAAAGCCAGT 2367  
QY 795 ATCTGTACTTACAGAGATCCAGTTTTCATGTCGAATTTCAAAGGCGATTCATTCAC 854  
DB 2366 ATCTGTACTTACAGAGATCCAGTTTTCATGTCGAATTTCAAAGGCGATTCATTCAC 2307  
QY 855 TACGAATGATGCAATTAACCACTCTGCTGTGTAATTTGACATTTCAATACAGACTT 914  
DB 2306 TACGAATGATGCAATTAACCACTCTGCTGTGTAATTTGACATTTCAATACAGACTT 2247  
QY 915 TTCTTATCAGCTTGAGATGCTTCAAGCTGATCTGCTTAAACGTGATTTGAGGTACA 974  
DB 2246 TTCTTATCAGCTTGAGATGCTTCAAGCTGATCTGCTTAAACGTGATTTGAGGTACA 2187  
QY 975 AAGCTTACTCCAAAGATGACAGCTTGAAGATTAAGAGAGCACTGCTCTTTGAAAT 1034  
DB 2186 AAGCTTACTCCAAAGATGACAGCTTGAAGATTAAGAGAGCACTGCTCTTTGAAAT 2127  
QY 1035 AAGGAGAGACAAAGAAAGAGAGCTTACCCAGCAATATCCTGCGGAGATGTT 1094  
DB 2126 AAGGAGAGACAAAGAAAGAGAGCTTACCCAGCAATATCCTGCGGAGATGTT 2067  
QY 1095 TCTCCAGTTCATTTTACCTGTGTCTTGAATTCGAGCAATTCCTAAAGGCAATTTT 1154  
DB 2066 TCTCCAGTTCATTTTACCTGTGTCTTGAATTCGAGCAATTCCTAAAGGCAATTTT 2007  
QY 1155 GCGAGGCTTGTGAGATATACAGTGAACGTCTGTAAGAAAGCGAGCTACAGAGCTG 1214  
DB 2006 GCGAGGCTTGTGAGATATACAGTGAACGTCTGTAAGAAAGCGAGCTACAGAGCTG 1947  
QY 1215 CAGTAAACAAGGGAGAGCCGATTAATAGCCGCTTGTAGAGATGCTGAGCTGTGTT 1274  
DB 1946 CAGTAAACAAGGGAGAGCCGATTAATAGCCGCTTGTAGAGATGCTGAGCTGTGTT 1887  
QY 1275 GATCTCTCTCTGCTGCTTCCCTTCTTGCAGGACACACACTCAGTCTCTCTGAAACAT 1334  
DB 1886 GATCTCTCTCTGCTGCTTCCCTTCTTGCAGGACACACACTCAGTCTCTCTGAAACAT 1827  
QY 1335 TCTTAAACTTGAACCCAGACCAATATTCGTGCAAGCTCAAGTTTATTCACCCAGAAA 1394  
DB 1826 TCTTAAACTTGAACCCAGACCAATATTCGTGCAAGCTCAAGTTTATTCACCCAGAAA 1767  
QY 1395 GCTTCATTTTGTCTTCAACATTTGTGAATTTCTGTCTACTGCAACAAGAGTCTGCG 1454  
DB 1766 GCTTCATTTTGTCTTCAACATTTGTGAATTTCTGTCTACTGCAACAAGAGTCTGCG 1707  
QY 1455 GAAGGAGATATTAACAGCTGAGCTGCTGCTTGTGTTGCTTCAAGTCTTCAAGCAACAT 1514  
DB 1706 GAAGGAGATATTAACAGCTGAGCTGCTGCTTGTGTTGCTTCAAGTCTTCAAGCAACAT 1647  
QY 1515 ACATGATCCATGAAGACAGC-GGAAAGCCCTGAGCTCTTAAGATATCATCTCTCTC 1573  
DB 1646 ACATGATCCATGAAGACAGCAGGAGAAAGCCCTGAGCTCTTAAGATATCATCTCTCTC 1587  
QY 1574 GAACAACAATTTCTTCCACTTAACAAGAGACCCCTCAATCCCATTAATGAGTGGGTC 1633  
DB 1586 GAACAACAATTTCTTCCACTTAACAAGAGACCCCTCAATCCCATTAATGAGTGGGTC 1527  
QY 1634 CAGGAACCGGATAGCCCGTTATATGGTCTCTCAACATAGAGAAACTCAAGAAC 1693  
DB 1526 CAGGAACCGGATAG-CCGTTTATTTGGTCTCTCAACATAGAGAAACTCAAGAAC 1468  
QY 1694 AACACCCAGATGAATTTTGAAGCAATGAGTGTGTTTGGCTGACAGGCAATAGATTA 1753  
DB 1467 AACACCCAGATGAATTTTGAAGCAATGAGTGTGTTTGGCTGACAGGCAATAGATTA 1408

QY 1754 GGGATTATCTATTCAGAAAAAGCTCAGACATTTCTTAAAGCATGGGATCTTAATCATC 1813  
| | | | |  
DB 1407 GGGATTATCTATTCAGAAAAAGCTCAGATTTCTTAAAGCATGGGATCTTAATCATC 1348  
| | | | |  
QY 1814 TAAAGTTTCTCTCTCAAGAGATGCTCTGTTGGGGAGAGAAAGCCCGCAAGATATG 1873  
| | | | |  
DB 1347 TAAAGTTTCTCTCTCAAGAGATGCTCTGTTGGGGAGAGAAAGCCCGCAAGATATG 1288  
| | | | |  
QY 1874 TACAGACACATTCACAGCTTCATGGCAGAGAGTGGAGAAATCTCTCCACAGAGAACG 1933  
| | | | |  
DB 1287 TACAGACACATTCACAGCTTCATGGCAGAGAGTGGAGAAATCTCTCCACAGAGAACG 1228  
| | | | |  
QY 1934 GCCATATTTATGTGTGTGAGATGCAAAAGATATGCGCAAGATGATCATGATGCTTTG 1993  
| | | | |  
DB 1227 GCCATATTTATGTGTGTGAGATGCAAAAGATATGCGCAAGATGATCATGATGCTTTG 1168  
| | | | |  
QY 1994 TGCATAATTAATGAGAAAGAGTTGGAGTTGAAAACATAGAAAGCAATGAAACCTTGCCCA 2053  
| | | | |  
DB 1167 TGCATAATTAATGAGAAAGAGTTGGAGTTGAAAACATAGAAAGCAATGAAACCTTGCCCA 1108  
| | | | |  
QY 2054 CTTTAAAGAGAAAGAAAGCGTACCTTCAGGATATTTGGTCATTA 2097  
| | | | |  
DB 1107 CTTTAAAGAGAAAGAAAGCGTACCTTCAGGATATTTGGTCATTA 1064  
| | | | |  
RESULT 9  
BD077780 390 bp DNA linear PAT 27-AUG-2002  
LOCUS 5'EST of secretory protein in brain.  
DEFINITION BD077780  
ACCESSION BD077780.1 GI:22623383  
VERSION JP 2001512015-A/65.  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
TITLE Edwards,J.B.D.M., Duclet,A. and Lacroix,B.  
JOURNAL 5'EST of secretory protein in brain  
PATENT: JP 2001512015-A 65 21-AUG-2001;  
GENSER  
COMMENT OS Homo sapiens (human)  
PN JP 2001512015-A/65  
PD 21-AUG-2001  
PF 31-JUL-1998 JP 2000505293  
PI 01-AUG-1997 US 08/905223  
PI JEAN BAPTISTE DUMAS MILNE EDWARDS,AYMERIC DUCLENT,BRUNO PI  
LACROIX  
PC C12N15/09,C07K14/47,C12M1/00,C12P21/02,C12Q1/68,C12N15/00 CC  
Von Heljne matrix  
CC score 6.9  
CC seq SLSLSHSHSVSC/SN  
FH Key Location/Qualifiers  
FT sig\_peptide 289..357.  
Location/Qualifiers  
SOURCE 1..390  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 18.4%; Score 386; DB 6; Length 390;  
Best Local Similarity 100.0%; Pred. No. 5.7e-199;  
Matches 386; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 970 GTACAAAGCTTACTCAAAAGCTGCAAGTTGAAGTTAAAGAGCACTGCGTCTTTTG 1029  
| | | | |  
DB 3 GTACAAAGCTTACTCAAAAGCTGCAAGTTGAAGTTAAAGAGCACTGCGTCTTTTG 62  
| | | | |  
QY 1030 AAAATTAAGGAGACACAAAGAGAGAGTACCTTACCCGACATATACCTGCGGGA 1089  
| | | | |  
DB 63 AAAATTAAGGAGACACAAAGAGAGAGTACCTTACCCGACATATACCTGCGGGA 122  
| | | | |

QY 1090 TGTCTCTCAGTTCATTTTACCTGTGTCTTGAATTCGAGCAATTCCTAAAAGGCA 1149  
| | | | |  
DB 123 TGTCTCTCAGTTCATTTTACCTGTGTCTTGAATTCGAGCAATTCCTAAAAGGCA 182  
| | | | |  
QY 1150 TTTTTCGAGCCCTTGTGACATTAACAGTGAACAGTGTGAAAAGGCGGCTACAGGAG 1209  
| | | | |  
DB 183 TTTTTCGAGCCCTTGTGACATTAACAGTGAACAGTGTGAAAAGGCGGCTACAGGAG 242  
| | | | |  
QY 1210 CTGTGACATTAACAGAGGCGAGCGGATTAATAGCCGTTTGTAGAGATGCTGTGCTGC 1269  
| | | | |  
DB 243 CTGTGACATTAACAGAGGCGAGCGGATTAATAGCCGTTTGTAGAGATGCTGTGCTGC 302  
| | | | |  
QY 1270 TTGTGATCTCTCTCTGCTTCTCTTCTTGGCCAGCCACTCATGCTCTGCTGAA 1329  
| | | | |  
DB 303 TTGTGATCTCTCTCTGCTTCTCTTCTTGGCCAGCCACTCATGCTCTGCTGAA 362  
| | | | |  
QY 1330 CATCTTCCTAACTTCAACCCAGACC 1355  
| | | | |  
DB 363 CATCTTCCTAACTTCAACCCAGACC 388  
| | | | |  
RESULT 10  
F121202804 1353 bp DNA linear PRI 14-DEC-1999  
LOCUS F121202804  
DEFINITION Homo sapiens methionine synthase reductase (WTRR) gene, exon 5.  
ACCESSION AF121205  
VERSION AF121205.1 GI:6572530  
KEYWORDS  
SEGMENT 4 of 12  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
TITLE Edwards,J.B.D.M., Wu,Q., Wilson,A., Hutzenga,J.J., Rozen,R.,  
JOURNAL Scherer,S.W. and Gravel,R.A.  
MEDLINE Molecular cloning, expression and physical mapping of the human  
Gene 240 (1), 75-88 (1999)  
PUBMED 20033550  
10564814  
REFERENCE 2 (bases 1 to 1353)  
AUTHORS Leclerc,D.  
TITLE Direct Submision  
JOURNAL Submitted (20-JAN-1999) Human Genetics, Montreal Children's  
Hospital, 4060 Ste-Catherine West, Montreal, Quebec H3Z 2Z3, Canada  
FEATURES  
SOURCE 1..1353  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/map="5p15.2-p15.3; between W11755 and D5S1957"  
exon 358..736  
/gene="WTRR"  
/number=5  
ORIGIN  
Query Match 18.2%; Score 381; DB 9; Length 1353;  
Best Local Similarity 100.0%; Pred. No. 2.8e-196;  
Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 401 GTTTAGAACTTGTGTTGACCGGTGATTTCTGGAATCTGGCAGCCCTCAGAAAGCATT 460  
| | | | |  
DB 357 GTTTAGAACTTGTGTTGACCGGTGATTTCTGGAATCTGGCAGCCCTCAGAAAGCATT 416  
| | | | |  
QY 461 TTAGGTCAACAGAGACAAAGAGATTAAGTGGCGCATCTCCGCTGGCATCACCTGCAT 520  
| | | | |  
DB 417 TTAGGTCAACAGAGACAAAGAGATTAAGTGGCGCATCTCCGCTGGCATCACCTGCAT 476  
| | | | |  
QY 521 CTTTGAAGAGACCTTGTGATTAAGTCAAGCTGTACACATTTGAATCTCAAGTGAAGCTTC 580  
| | | | |

Db 477 CCTTGAAGACAGACCTTGTGAAGTCAAGAGCTGCTACACATTGAATCTCAAGTCAGCTTC 536  
QY 581 TGAATTCATGATTCAGAGAAAGAAAGATTCAGGTTTGAAGCAAAATGACAGTAACA 640  
Db 537 TGAATTCATGATTCAGAGAAAGAAAGATTCAGGTTTGAAGCAAAATGACAGTAACA 596  
QY 641 GCAACCAATCCAAATGTTGAATTAATGAAGACTTGAAGTCTCACTTACCCGTTGGTACCCC 700  
Db 597 GCAACCAATCCAAATGTTGAATTAATGAAGACTTGAAGTCTCACTTACCCGTTGGTACCCC 656  
QY 701 CACTCTCACAAGCCTCTCTGAATATTCCTGTTTACCCCAAGAAATTTACAGTATC 760  
Db 657 CACTCTCACAAGCCTCTCTGAATATTCCTGTTTACCCCAAGAAATTTACAGTATC 716  
QY 761 TGCAGAGTCTCTTGGCCAGG 781  
Db 717 TGCAGAGTCTCTTGGCCAGG 737

RESULT 11  
AC010346 109626 bp DNA linear PRI 10-NOV-2000  
LOCUS Homo sapiens chromosome 5 clone C17B-H1\_2018B2, complete sequence.  
AC010346 GI:11136705  
DEFINITION HTG.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM

REFERENCE  
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
TITLE 1 (bases 1 to 109626)  
REFERENCE DOB Joint Genome Institute and Stanford Human Genome Center.  
AUTHORS Unpublished  
TITLE 2 (bases 1 to 109626)  
REFERENCE DOB Joint Genome Institute.  
AUTHORS Direct Submission  
TITLE Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint  
JOURNAL Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
3 (bases 1 to 109626)  
REFERENCE DOB Joint Genome Institute and Stanford Human Genome Center.  
AUTHORS Direct Submission  
TITLE Submitted (10-NOV-2000) DOE Joint Genome Institute, 2800 Mitchell  
JOURNAL Drive, Walnut Creek, CA 94598, USA  
On Nov 10, 2000 this sequence version replaced gi:9256196.  
COMMENT Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
www-shgc.stanford.edu  
Quality: Phrap Quality >=40 99.9% of Sequence;  
SFS Content:  
WI-9255 G05749.

## FEATURES

source  
1. 109626  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/clone="C17B-H1\_2018B2"

## ORIGIN

Query Match 15.7%; Score 330; DB 9; Length 109626;  
Best Local Similarity 99.7%; Pred. No. 1.6e-168; Indels 0; Gaps 0;  
Matches 380; Conservative 0; Mismatches 1;  
QY 401 GTTTAGAACTTGTGTGAGCCGTGATTCCTGAGCTCTGCGCAGCCCTCAGAAAGCAT 460  
Db 88571 GTTTAGAACTTGTGTGAGCCGTGATTCCTGAGCTCTGCGCAGCCCTCAGAAAGCAT 88630  
QY 461 TTTGGTCAAGCAGAGCAGAGAGATTAAGTGGCGCACTCCCGGTGGCATCACTGGCAT 520  
Db 88631 TTTGGTCAAGCAGAGCAGAGAGATTAAGTGGCGCACTCCCGGTGGCATCACTGGCAT 88690

QY 521 CCTTGAAGACAGACCTTGTGAAGTCAAGAGCTGCTACACATTGAATCTCAAGTCAGCTTC 580  
Db 88691 CCTTGAAGACAGACCTTGTGAAGTCAAGAGCTGCTACACATTGAATCTCAAGTCAGCTTC 88750  
QY 581 TGAATTCATGATTCAGAGAAAGAAAGATTCAGGTTTGAAGCAAAATGACAGTAACA 640  
Db 88751 TGAATTCATGATTCAGAGAAAGAAAGATTCAGGTTTGAAGCAAAATGACAGTAACA 88810  
QY 641 GCAACCAATCCAAATGTTGAATTAATGAAGACTTGAAGTCTCACTTACCCGTTGGTACCCC 700  
Db 88811 GCAACCAATCCAAATGTTGAATTAATGAAGACTTGAAGTCTCACTTACCCGTTGGTACCCC 88870  
QY 701 CACTCTCACAAGCCTCTCTGAATATTCCTGTTTACCCCAAGAAATTTACAGTATC 760  
Db 88871 CACTCTCACAAGCCTCTCTGAATATTCCTGTTTACCCCAAGAAATTTACAGTATC 88930  
QY 761 TGCAGAGTCTCTTGGCCAGG 781  
Db 88931 TGCAGAGTCTCTTGGCCAGG 88951

RESULT 12  
AC025174 110756 bp DNA linear PRI 28-MAR-2002  
LOCUS Homo sapiens chromosome 5 clone CTD-2072124, complete sequence.  
AC025174 GI:19774456  
DEFINITION HTG.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM

REFERENCE  
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
TITLE 1 (bases 1 to 110756)  
REFERENCE DOB Joint Genome Institute and Stanford Human Genome Center.  
AUTHORS Unpublished  
TITLE 2 (bases 1 to 110756)  
REFERENCE DOB Joint Genome Institute.  
AUTHORS Direct Submission  
TITLE Submitted (07-MAR-2002) Production Sequencing Facility, DOE Joint  
JOURNAL Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
3 (bases 1 to 110756)  
REFERENCE DOB Joint Genome Institute.  
AUTHORS Direct Submission  
TITLE Submitted (07-MAR-2002) Production Sequencing Facility, DOE Joint  
JOURNAL Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
4 (bases 1 to 110756)  
REFERENCE DOB Joint Genome Institute and Stanford Human Genome Center.  
AUTHORS Direct Submission  
TITLE Submitted (28-MAR-2002) DOE Joint Genome Institute, 2800 Mitchell  
JOURNAL Drive, Walnut Creek, CA 94598, USA  
On Mar 28, 2002 this sequence version replaced gi:19224767.  
COMMENT Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
www-shgc.stanford.edu  
Quality: Phrap Quality >=40 100% of Sequence;  
Estimated Total Number of Errors is 0.

## FEATURES

source  
1. 110756  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/clone="CTD-2072124"

## ORIGIN

Query Match 15.7%; Score 330; DB 9; Length 110756;  
Best Local Similarity 99.7%; Pred. No. 1.6e-168; Indels 0; Gaps 0;  
Matches 380; Conservative 0; Mismatches 1;  
QY 401 GTTTAGAACTTGTGTGAGCCGTGATTCCTGAGCTCTGCGCAGCCCTCAGAAAGCAT 460



Db 20040 GTTTAGACTTGTGGTTGAGCCGTGACTTCTGAGCTTGGCCAGCCCTCAGAAAGCATT 20099  
QY 461 TTAGTCACAGAGAGCAAGAGAGAGTAAGTGGCGCACTCCGGTGGCATCACTGCAT 520  
Db 20100 TTAGTCACAGAGAGCAAGAGAGAGTAAGTGGCGCACTCCGGTGGCATCACTGCAT 20159  
QY 521 CCTTGAAGACAGACCTTGTGAAGTCAAGTCTGCTACACATTGAATCTCAAGTGGAGCTTC 580  
Db 20160 CCTCGAGAGACAGACCTTGTGAAGTCAAGTCTGCTACACATTGAATCTCAAGTGGAGCTTC 20219  
QY 581 TGAATTCGATGATTCAGAAAGAGATTCGAGGTTTGAAGCAAAAGCAGTACACA 640  
Db 20220 TGAATTCGATGATTCAGAAAGAGATTCGAGGTTTGAAGCAAAAGCAGTACACA 20279  
QY 641 GCAACCAATCAATGTTGTAATTAAGAGACTTGAAGTCTCACTTACCCGTTGGTACCCC 700  
Db 20280 GCAACCAATCAATGTTGTAATTAAGAGACTTGAAGTCTCACTTACCCGTTGGTACCCC 20339  
QY 701 CACTCTCACAAGCCTCTCTGAATTTCTCGTGTAACTCCCAAGAAATTTACAGGTACATC 760  
Db 20340 CACTCTCACAAGCCTCTCTGAATTTCTCGTGTAACTCCCAAGAAATTTACAGGTACATC 20399  
QY 761 TGCAGAGTCTCTTGGCCAGG 781  
Db 20400 TGCAGAGTCTCTTGGCCAGG 20420

RESULT 13  
AC022921  
LOCUS  
DEFINITION Homo sapiens clone RP11-138P20, WORKING DRAFT SEQUENCE, 12  
unordered pieces.  
AC022921  
AC022921.2 GI:7229868  
VERSION  
KEYWORDS HTG; HTGS PHASE1; HTGS\_DRAFT.  
SOURCE  
ORGANISM Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 158199)  
Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
Homo sapiens, clone RP11-138P20  
Unpublished  
2 (bases 1 to 158199)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,  
Bogunlavsky,L., Bouhgalter,B., Brown,A., Burkett,R., Castle,A.,  
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,  
Dearellano,K., Dewar,K., Domino,M., Doyle,M., Fenesfor,J.,  
Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J.,  
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,  
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,  
Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,  
Macdonald,P., Margis,N., McEwan,P., McGuck,A., McKernan,K.,  
McPheters,R., Meldrum,J., Meneus,L., Morrow,J., Naylor,J.,  
Norman,C.H., O'Connor,T., O'Donnell,P., Olivari,T.M., Peterson,K.,  
Pierre,N., Pisanl,C., Pollara,V., Raymond,C., Riley,R., Rochman,D.,  
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,  
Stojanovic,N., Subramanian,A., Talamas,J., Teefay,S., Theodore,J.,  
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.D.,  
Zimmer,A. and Zody,M.

TITLE  
JOURNAL  
COMMENT  
Direct Submission  
Submitted (07-FEB-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Mar 12, 2000 this sequence version replaced gi:6921909.  
All repeats were identified using RepeatMasker:  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WITB  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information  
Center project name: L6314  
Center clone name: 138 P 20  
----- Summary Statistics  
Sequencing vector: M13; M77815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 152636 bases at least Q40  
Consensus quality: 155474 bases at least Q30  
Consensus quality: 156388 bases at least Q20  
Insert size: 178000; agarose-fp  
Insert size: 157099; sum-of-contigs  
Quality coverage: 4.4 in Q20 bases; agarose-fp  
Quality coverage: 5.0 in Q20 bases; sum-of-contigs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 12 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1 1283: contig of 1283 bp in length  
1284 1383: gap of 100 bp  
1384 4203: contig of 2820 bp in length  
4204 4303: gap of 100 bp  
4304 6786: contig of 2483 bp in length  
6787 6886: gap of 100 bp  
6887 9683: contig of 2797 bp in length  
9684 9783: gap of 100 bp  
9784 12902: contig of 3119 bp in length  
12903 13002: gap of 100 bp  
13003 16429: contig of 3427 bp in length  
16430 16529: gap of 100 bp  
16530 25201: contig of 8672 bp in length  
25202 25301: gap of 100 bp  
25302 36759: contig of 11458 bp in length  
36760 36859: gap of 100 bp  
36860 53921: contig of 17062 bp in length  
53922 54021: gap of 100 bp  
54022 72054: contig of 18033 bp in length  
72055 72154: gap of 100 bp  
72155 102527: contig of 30373 bp in length  
102528 102627: gap of 100 bp  
102628 158199: contig of 55572 bp in length.  
----- Location/Qualifiers  
1. 158199  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="RP11-138P20"  
/clone\_1lb="RP11-138P20"  
1. 1283  
/note="assembly\_fragment"  
1384. 4203  
/note="assembly\_fragment"  
4304. 6786  
/note="assembly\_fragment"  
6887. 9683  
/note="assembly\_fragment"  
9784. 12902  
vector\_side:left"  
vector\_side:right"  
13003. 16429  
/note="assembly\_fragment"  
16530. 25201  
/note="assembly\_fragment"  
clone\_end:T7  
vector\_side:right"  
25302. 36759  
/note="assembly\_fragment"



```

misc_feature      36860..53921
                  /note="assembly_fragment"
misc_feature      54022..72054
                  /note="assembly_fragment"
misc_feature      72155..102527
                  /note="assembly_fragment"
misc_feature      102628..158199
                  /note="assembly_fragment"

ORIGIN
Query Match      13.3%; Score 279; DB 2; Length 158199;
Best Local Similarity 99.5%; Pred. No. 1,3e-140;
Matches 379; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 401 GTTTAAGACTTGTGTTGAGCCGTGATTCCTGACTCTGGCCAGCCCTCAGAAAGCATT 460
DB 3446 GTTTAAGACTTGTGTTGAGCCGTGATTCCTGACTCTGGCCAGCCCTCAGAAAGCATT 3505
QY 461 TTAGGTCAAGCAGAGGCAAGAGAGATAGTGGCGCACTCCGGTGGCATCCTGCAT 520
DB 3506 TTAGGTCAAGCAGAGGCAAGAGAGATAGTGGCGCACTCCGGTGGCATCCTGCAT 3565
QY 521 CCTTGAAGCAGAGCCTTGTGAGTCTAGAGCTGCTACACATTGAATCTCAAGTCGAGCTTC 580
DB 3566 CCTTGAAGCAGAGCCTTGTGAGTCTAGAGCTGCTACACATTGAATCTCAAGTCGAGCTTC 3625
QY 581 TGAAGTTCGATGATTCAGAGAAAGAGATTCGAGGTTTGAAGCAAAATGCAAGTGAACA 640
DB 3626 TGAAGTTCGATGATTCAGAGAAAGAGATTCGAGGTTTGAAGCAAAATGCAAGTGAACA 3685
QY 641 GCACCAATCCATGTTGTAATGAAGACTTGAATCCTCACTACCCCGTTGGTACCCC 700
DB 3686 GCACCAATCCATGTTGTAATGAAGACTTGAATCCTCACTACCCCGTTGGTACCCC 3745
QY 701 CACTCTCAGACGCTCTCTGATATTCCTGTTTACCCCGAGATATTTACAGATACATC 760
DB 3746 CACTCTCAGACGCTCTCTGATATTCCTGTTTACCCCGAGATATTTACAGATACATC 3805
QY 761 TGCAGAGTCTCTTGGCCAGG 781
DB 3806 TGCAGAGTCTCTTGGCCAGG 3826

RESULT 14
AC021609      167237 bp DNA linear HTG 12-MAR-2000
LOCUS      Homo sapiens clone RP11-259D10, WORKING DRAFT SEQUENCE, 6 unordered
DEFINITION      pieces.
ACCESSION      AC021609
VERSION      AC021609.3 GI:7230210
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 167237)
AUTHORS      Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
               Boguslavskiy,L., Boukhalter,B., Brown,A., Burkett,G., Castle,A.,
               Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
               DeArrellano,K., Dewar,K., Domingo,M., Doyle,M., Fennestor,J.,
               Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
               Gardyna,S., Grant,G., Hagos,B., Heathford,A., Horton,L.,
               Howland,J.C., Johnson,R., Jones,C., Kam,L., Karasik,A., Klein,J.,
               Landers,T., Lenoczky,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
               Macdonald,P., Margulis,N., McEwan,P., McGurk,A., McKernan,K.,
               McPheeters,R., Meldrim,J., Menus,L., Morrow,J., Naylor,J.,
               Norman,C.H., O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K.,
               Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,

```

```

TITLE
JOURNAL
COMMENT
Roy,A., Santoe,R., Severy,P., Spencer,B., Strange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A., and Zody,M.
Direct Submission
Submitted (16-JUN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 12, 2000 this sequence version replaced gi:689697.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green,P. (1996-1997)
http://ftp.genome.washington.edu/XM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seg.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: 259 D 10
Center clone name: 259 D 10
----- Summary Statistics
Sequencing vector: M13; M7815, 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 16183 bases at least Q40
Consensus quality: 164380 bases at least Q30
Consensus quality: 165590 bases at least Q20
Insert size: 164000; agarose-fp
Insert size: 166737; sum-of-ctnigs
Quality coverage: 5.1 in Q20 bases; sum-of-ctnigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 3656: contig of 3656 bp in length
* 3657 3756: gap of 100 bp
* 3757 9436: contig of 5680 bp in length
* 9437 9536: gap of 100 bp
* 9537 27768: contig of 18222 bp in length
* 27769 27868: gap of 100 bp
* 27869 52058: contig of 24190 bp in length
* 52059 80100: contig of 27942 bp in length
* 80101 80200: gap of 100 bp
* 80201 167237: contig of 87037 bp in length.
----- Location/Qualifiers
1. 167237
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="RP11-259D10"
/clone_lib="RP11 Human Male BAC"
1. 3656
/note="assembly_fragment"
clone_end:17
vector_side:left"
misc_feature 3757..9436
/note="assembly_fragment"
misc_feature 9537..27768
/note="assembly_fragment"
clone_end:SP6
vector_side:right"
misc_feature 27869..52058
/note="assembly_fragment"
misc_feature 52159..80100
/note="assembly_fragment"
80201..167237
/note="assembly_fragment"

```



```
Matches 379; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 401 GTTTAGAACTTGTGTTGAGCCGTGATGCTGGACTTGCCAGCCCTCAGAAAGCATT 460
    |||||
Db 83524 GTTTAGAACTTGTGTTGAGCCGTGATGCTGGACTTGCCAGCCCTCAGAAAGCATT 83583
QY 461 TTAGGTCAAGCAGAGGACAGAGAGATAGTGGCGCACTCCCGGTGGCATCACTGCAT 520
    |||||
Db 83584 TTAGGTCAAGCAGAGGACAGAGAGATAGTGGCGCACTCCCGGTGGCATCACTGCAT 83643
QY 521 CCTTGAGGACAGACCTTGTGAAGTCAAGCTGCTACACATGAACTTCAAGTGCAGCTTC 580
    |||||
Db 83644 CCTTGAGGACAGACCTTGTGAAGTCAAGCTGCTACACATGAACTTCAAGTGCAGCTTC 83703
QY 581 TGAGATTGATGATTCAAGAAAGAAAGATTCTGAGGTTTGAAGCAAAATGACGTGACA 640
    |||||
Db 83704 TGAGATTGATGATTCAAGAAAGAAAGATTCTGAGGTTTGAAGCAAAATGACGTGACA 83763
QY 641 GCAACCAATCCAATGTGTAAATTGAAGACTTTGAGTCTCACTTACCCGTTGCGTACCC 700
    |||||
Db 83764 GCAACCAATCCAATGTGTAAATTGAAGACTTTGAGTCTCACTTACCCGTTGCGTACCC 83823
QY 701 CACTCTCAAGGCTCTCTGAATATTCCTGTTTACCCCGAATATTTACAGGTACATC 760
    |||||
Db 83824 CACTCTCAAGGCTCTCTGAATATTCCTGTTTACCCCGAATATTTACAGGTACATC 83883
QY 761 TGCAGAGAGTCTCTTGSCCAGG 781
    |||||
Db 83884 TGCAGAGAGTCTCTTGSCCAGG 83904
```

Search completed: August 27, 2005, 09:38:53  
Job time : 6011.21 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 13:32:20 ; Search time 733.49 Seconds  
(without alignment)  
16924.161 Million cell updates/sec

Title: US-09-371-347A-41

Perfect score: 2097  
Sequence: 1 atgagggaggttcctact.....ctcagatattgtgtcctaa 2097

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapept 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8760412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq16Dec04:\*  
1: geneseqn19808:\*  
2: geneseqn19908:\*  
3: geneseqn20008:\*  
4: geneseqn20018:\*  
5: geneseqn20028:\*  
6: geneseqn20038:\*  
7: geneseqn20048:\*  
8: geneseqn20058:\*  
9: geneseqn20068:\*  
10: geneseqn20078:\*  
11: geneseqn20088:\*  
12: geneseqn20098:\*  
13: geneseqn20108:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2094	99.9	2094	11	ADM43209 Human met
2	2046	97.6	3259	5	AA65070 DNA encod
3	2046	97.6	3259	5	AA65070 DNA encod
4	2046	97.6	3259	5	AA65070 DNA encod
5	2043	97.4	2094	11	ADM43208 Human ful
6	1995	95.1	3270	13	ADM43208 Human ful
7	1992	95.0	2094	11	ADM43212 Human met
8	1944	92.7	3259	3	AA58935 DNA encod
9	1800	85.8	2091	11	ADM43216 Human met
10	1800	85.8	2091	11	ADM43216 Human met
11	1701	81.1	3256	3	AA58977 A human m
12	1640	78.2	3255	3	AA58976 A human m
13	1062	50.6	3256	13	ADM39029 Human SNP
14	1062	50.6	3274	13	ADM39030 Human SNP
15	1007	48.0	3189	4	AA42470 Human dia
16	805	38.4	1986	4	AA41064 CDNA enco
17	427	20.4	1663	4	AA41602 CDNA enco
18	386	18.4	390	2	AA51820 Human sec
19	330	15.7	591	12	ACH73174 Human gen
20	328	15.6	379	12	ACH86905 Human gen

21	279	13.3	591	12	ACH868540	ACH66540 Human gen
22	277	13.2	379	12	ACH82240	ACH82240 Human gen
23	225	10.7	503	5	AA65069	AA65069 DNA encod
24	188	9.0	525	12	ACH67438	ACH67438 Human gen
25	175	8.3	175	12	ACH81143	ACH81143 Human gen
26	158	7.5	2475	6	AD32365	AD32365 Human jun
27	158	7.5	2475	13	AD161720	AD161720 Human gen
28	137	6.5	525	12	ACH73117	ACH73117 Human gen
29	124	5.9	175	12	ACH68848	ACH68848 Human gen
30	65	3.1	244	3	AA242736	AA242736 Human 5'
31	60	2.9	60	6	ABN36264	ABN36264 Human ep1
32	52	2.5	1835	5	AA65071	AA65071 DNA encod
33	51	2.4	51	4	AA178548	AA178548 Human 811
34	30	1.4	1681	11	AD131127	AD131127 Human CDN
35	26	1.2	26	3	AA58955	AA58955 PCR prime
36	26	1.2	26	3	AA58939	AA58939 PCR prime
37	26	1.2	26	6	ABX09549	ABX09549 Arteriosc
38	26	1.2	26	6	AA143713	AA143713 Ptegestat
39	26	1.2	26	11	ADM43205	ADM43205 Human met
40	26	1.2	26	11	ADM43189	ADM43189 Human met
41	25	1.2	25	3	AA58952	AA58952 PCR prime
42	25	1.2	25	3	AA58937	AA58937 PCR prime
43	25	1.2	25	3	AA58947	AA58947 PCR prime
44	25	1.2	25	11	ADM43187	ADM43187 Human met
45	25	1.2	25	11	ADM43202	ADM43202 Human met

## ALIGNMENTS

RESULT 1	ADM43209	ADM43209 standard; CDNA; 2094 BP.
ID	ADM43209	ADM43209 standard; CDNA; 2094 BP.
XX	ADM43209;	
AC	ADM43209;	
XX	03-JUN-2004 (first entry)	
DT	03-JUN-2004 (first entry)	
XX	Human methionine synthase reductase CDS G66A variant.	
XX	Human; sb; Methionine synthase reductase polypeptide; HsMTRR; cancer;	
KW	cardiovascular disease; neural tube defect; hyperhomocysteinemia;	
KW	chromosome 5p15.2-p15.3; SNP; single nucleotide polymorphism.	
XX		
OS	Homo sapiens.	
XX		
FX	Key	Location/Qualifiers
FT	CDS	1..2094
FT		/*tag= a
FT		//product= "HsMTRR"
FT		//partial
FT	variation	/note= "No stop codon shown"
FT		replace(66,G)
FT		/*tag= b
FT	variation	/standard name= "Single_nucleotide_polymorphism"
FT		replace(110,A)
FT		/*tag= c
FT		/standard_name= "Single_nucleotide_polymorphism"
PN	US2003082676-A1.	
XX		
PD	01-MAY-2003.	
XX		
XX	10-AUG-1999;	99US-00371347.
PF		
XX	16-JAN-1998;	98US-0071622P.
PR		
XX	15-JAN-1999;	99US-00232028.
PA	(GRAV/) GRAVEL R. A.	
PA	(ROZE/) ROZEN R.	
PA	(LECL/) LECLERC D.	
PA	(WILS/) WILSON A.	
PA	(ROSE/) ROSENBLATT D.	

XX Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;  
PI WPI; 2003-576610/54.  
XX P-PSDB; ADM43211.  
XX New substantially pure nucleic acid encoding a mammalian methionine  
PT synthase reductase polypeptide, useful for diagnosing, preventing or  
PT treating conditions associated with altered methionine synthase activity,  
e.g. cancer.

Claim 3; SEQ ID NO 41; 26bp; English.

XX The invention relates to a substantially pure nucleic acid that encodes a  
XX mammalian methionine synthase reductase polypeptide, HemTR, or that  
XX hybridizes at high stringency to a nucleic acid appearing as ADM43208 or  
XX ADM43209. Also included are a non-human animal where one or both genetic  
XX alleles encoding the methionine synthase reductase polypeptide are  
XX mutated, an antibody that specifically binds the above methionine  
XX synthase reductase polypeptide, a method of detecting the presence of the  
XX above polypeptide, a method for detecting sequence variants for  
XX methionine synthase reductase in a mammal, methods of treating or  
XX preventing cancer (or cardiovascular disease or neural tube defects) in a  
XX subject, methods of screening for a compound that modulates methionine  
XX synthase reductase biological activity and a method for detecting an  
XX increased risk of developing a neural tube defect in a mammalian embryo  
XX or foetus. The nucleic acid is useful in diagnosing, preventing or  
XX treating conditions associated with altered methionine synthase activity,  
XX such as cancer, cardiovascular disease or neural tube defects, or in  
XX screening for a compound that modulates methionine synthase reductase  
XX biological activity. Naturally occurring variants of the polypeptide are  
XX also associated with hyperhomocysteinemia. The gene for HemTR is  
XX located on chromosome 5p15.2-p15.3. The present sequence is the coding  
XX sequence of a variant human hemTR cDNA.

Sequence 2094 BP; 592 A; 489 C; 480 G; 533 T; 0 U; 0 Other;

Query Match 99.9%; Score 2094; DB 11; Length 2094;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2094; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGAGAGGTTCTGTACTATATGCTACACAGAGGAGGAGGCAAGGCAATCGAGAA 60  
DB 1 ATGAGAGAGGTTCTGTACTATATGCTACACAGAGGAGGAGGCAAGGCAATCGAGAA 60  
QY 61 GAAATATGTAGCAGACGTGTGTACATGATTTTCTGCAGATCTTCACTGATTAAGTGA 120  
DB 61 GAAATATGTAGCAGACGTGTGTACATGATTTTCTGCAGATCTTCACTGATTAAGTGA 120  
QY 121 TCCGATTAAGTATGACCTTAACAAACCGAAACAGCTCCTCTTGTGTTGTGTTCTACACG 180  
DB 121 TCCGATTAAGTATGACCTTAACAAACCGAAACAGCTCCTCTTGTGTTGTGTTCTACACG 180  
QY 181 GGACACCGAGAGACCCAGACAGAGCCGAGTTGTTAAGAAATPAGAAACCAAA 240  
DB 181 GGACACCGAGAGACCCAGACAGAGCCGAGTTGTTAAGAAATPAGAAACCAAA 240  
QY 241 CTGCCGCTGATTTCTTTGTCTCACTGCGGTATGGTTACTGGGTCCTCGGTATTCAGAA 300  
DB 241 CTGCCGCTGATTTCTTTGTCTCACTGCGGTATGGTTACTGGGTCCTCGGTATTCAGAA 300  
QY 301 TACACCTACTTTTGGCAATGGGGGGAAGATTAATTGAATTAACGACTTCAAGAGCTTGAAGCC 360  
DB 301 TACACCTACTTTTGGCAATGGGGGGAAGATTAATTGAATTAACGACTTCAAGAGCTTGAAGCC 360  
QY 361 CGGCATTTCATATGACCTGGAGCATGAGATGACTGTGTAGTTTAAACCTGTGGTTGAG 420  
DB 361 CGGCATTTCATATGACCTGGAGCATGAGATGACTGTGTAGTTTAAACCTGTGGTTGAG 420  
QY 421 CCGTGAATGCTGAGCTCTGGCCAGGCTTCAGAAACATTTTAAAGTCAAGAGAGAGCA 480  
DB 421 CCGTGAATGCTGAGCTCTGGCCAGGCTTCAGAAACATTTTAAAGTCAAGAGAGAGCA 480

QY 481 GAGGAGATTAAGTGGCGCACTCCCGGTGGCATATCACTGCATCTTTAGAGACAGACTTGTG 540  
DB 481 GAGGAGATTAAGTGGCGCACTCCCGGTGGCATATCACTGCATCTTTAGAGACAGACTTGTG 540  
QY 541 AAGTCAGAGCTCTACATCAATCTCAAGTCAGAGCTTCTGAGATTCGATGATTGAGGA 600  
DB 541 AAGTCAGAGCTCTACATCAATCTCAAGTCAGAGCTTCTGAGATTCGATGATTGAGGA 600  
QY 601 AGAAAGATTCGAGATTTTGAAGCAAAATGCAATGCAAGCAACCAATCCAAATGTTGA 660  
DB 601 AGAAAGATTCGAGATTTTGAAGCAAAATGCAATGCAAGCAACCAATCCAAATGTTGA 660  
QY 661 ATTGAAGACTTGAATCCTCACTACCTGCTGGTATCCCGCACTCTCAACAGCTCTCTG 720  
DB 661 ATTGAAGACTTGAATCCTCACTACCTGCTGGTATCCCGCACTCTCTCAACAGCTCTCTG 720  
QY 721 AATATTCCTGGTTTACCCCGCAAAATTTTACAGATCACTGAGAGATCTTGGCCAG 780  
DB 721 AATATTCCTGGTTTACCCCGCAAAATTTTACAGATCACTGAGAGATCTTGGCCAG 780  
QY 781 GAGGAAAGCAAGTATCTGTGACTTCAGAGATCCAGTTTTCAGATGCAATTTCAAG 840  
DB 781 GAGGAAAGCAAGTATCTGTGACTTCAGAGATCCAGTTTTCAGATGCAATTTCAAG 840  
QY 841 GCATTTCAACTTACTACGATATGATCCATTAACCACTCTGCTGTGATTTGACAT 900  
DB 841 GCATTTCAACTTACTACGATATGATCCATTAACCACTCTGCTGTGATTTGACAT 900  
QY 901 TCAATTAACAGACTTTTCTATCAGCTGAGATGAGTCTGAGGATCTGAGCTTAAAGT 960  
DB 901 TCAATTAACAGACTTTTCTATCAGCTGAGATGAGTCTGAGGATCTGAGCTTAAAGT 960  
QY 961 GATTTGAGATCAAAAGCTTACCTCAAGATCTGAGCTTGAAGATTAAGAGAGAGAGCTGC 1020  
DB 961 GATTTGAGATCAAAAGCTTACCTCAAGATCTGAGCTTGAAGATTAAGAGAGAGAGCTGC 1020  
QY 1021 GTCTTTTGAATAATTAAGGACACACAAAGAAAGAGAGAGCTTACCTCAAGATTA 1080  
DB 1021 GTCTTTTGAATAATTAAGGACACACAAAGAAAGAGAGAGCTTACCTCAAGATTA 1080  
QY 1081 CCTGGGGAGTGTCTCTCAAGTTCAATTTTACCTGCTGCTTGAATCCAGACATTTCT 1140  
DB 1081 CCTGGGGAGTGTCTCTCAAGTTCAATTTTACCTGCTGCTTGAATCCAGACATTTCT 1140  
QY 1141 AAAAAAGCATTTTTCGAGCCCTTGTGACATTAACAGTGAAGTGTGAAAAAGGCGAG 1200  
DB 1141 AAAAAAGCATTTTTCGAGCCCTTGTGACATTAACAGTGAAGTGTGAAAAAGGCGAG 1200  
QY 1201 CTACAGAGCTGTGAGATTAACAGAGGCGAGCCGATTAAGCCGCTTTGTAGAGATGCG 1260  
DB 1201 CTACAGAGCTGTGAGATTAACAGAGGCGAGCCGATTAAGCCGCTTTGTAGAGATGCG 1260  
QY 1261 TGTGCTGCTGTGTGATCTCTCTGCTGCTTCCCTCTGTGAGAGCACTCAAGTCTC 1320  
DB 1261 TGTGCTGCTGTGTGATCTCTCTGCTGCTTCCCTCTGTGAGAGCACTCAAGTCTC 1320  
QY 1321 CTGCTCGAATCTTCTCTAACTTCAACCCAGACCAATATTCGTGTGCAAGCTCAAGTTA 1380  
DB 1321 CTGCTCGAATCTTCTCTAACTTCAACCCAGACCAATATTCGTGTGCAAGCTCAAGTTA 1380  
QY 1381 TTTTCAACCCAGAAAGCTCATTTTGTCTTCAAACTTGTGGAATTTCTGTCTACATGCGACA 1440  
DB 1381 TTTTCAACCCAGAAAGCTCATTTTGTCTTCAAACTTGTGGAATTTCTGTCTACATGCGACA 1440  
QY 1441 ACAGAGTTCTGCGAGAGAGATATGATCAAGCTGAGCTGAGCTTGTGATTTGCTTCAAGTT 1500  
DB 1441 ACAGAGTTCTGCGAGAGAGATATGATCAAGCTGAGCTGAGCTTGTGATTTGCTTCAAGTT 1500  
QY 1501 CTTTACGCAAAATTAATGATCCCATGAAGACAGCGGAAAAGCCCTGCTCTTAAGATA 1560  
DB 1501 CTTTACGCAAAATTAATGATCCCATGAAGACAGCGGAAAAGCCCTGCTCTTAAGATA 1560  
QY 1561 TCCATCTCTCTCGAACAACAATTTCTTCCACTTAACAGATGACCCCTCAATCCCATC 1620

Db	1561	TTCAATCTCTCTCGAACAACAATTCCTTCCACTTACCAAGATGACCCCTCAATCCCATC	1620
Qy	1621	ATAATGCTGGTCCAGGAACCGGCATAGCCCGTTTATTGGGTTCTTCAACATATAGAG	1680
Db	1621	ATAATGCTGGTCCAGGAACCGGCATAGCCCGTTTATTGGGTTCTTCAACATATAGAG	1680
Qy	1681	AAATCTCAAGAACCAACCCGAGATGGAAATTTTGAGCAATGAGTGTTTTTTGGCTGC	1740
Db	1681	AAATCTCAAGAACCAACCCGAGATGGAAATTTTGAGCAATGAGTGTTTTTTGGCTGC	1740
Qy	1741	AGGCATTAAGATAGGGAATTATCTATTCAGAAAAGCTCAGACATTTCTTAAACATAGG	1800
Db	1741	AGGCATTAAGATAGGGAATTATCTATTCAGAAAAGCTCAGACATTTCTTAAACATAGG	1800
Qy	1801	ATCTTAACTCATCTPAAAGTTTCTTTCAGAGAGATGCTCTGTGTGGGAGAGAGAACCC	1860
Db	1801	ATCTTAACTCATCTPAAAGTTTCTTTCAGAGAGATGCTCTGTGTGGGAGAGAGAACCC	1860
Qy	1861	CCACCAAAAGTATGTATCAAGACAACATCCAGTTTCATATGCCAGACAGTGGCGAANAATCTC	1920
Db	1861	CCACCAAAAGTATGTATCAAGACAACATCCAGTTTCATATGCCAGACAGTGGCGAANAATCTC	1920
Qy	1921	CTCAGAGGAACGGCCATATTATTATGTGTGAGATGCAGAAAGATATATGGCCAAAGATGTA	1980
Db	1921	CTCAGAGGAACGGCCATATTATTATGTGTGAGATGCAGAAAGATATATGGCCAAAGATGTA	1980
Qy	1981	CATGATGCTCTTGTGCAATATATTAAGCAAAAGGTTGAGGTTGAAAACTGAAAGCAATG	2040
Db	1981	CATGATGCTCTTGTGCAATATATTAAGCAAAAGGTTGAGGTTGAAAACTGAAAGCAATG	2040
Qy	2041	AAAACCCCTGGCACTTTTAAAGAAAGAAAAGCTACCTTCAGAGATTTTGGTCA	2094
Db	2041	AAAACCCCTGGCACTTTTAAAGAAAGAAAAGCTACCTTCAGAGATTTTGGTCA	2094

## RESULT 2

ID AAS65070 standard; cDNA; 3259 BP.

AC AAS65070;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #874

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX

XX  
XX

XX

[illegible]

**XX**

PR 23-AUG-2000; 2000US-00649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

XX DE

PT diagnostics, forensics, gene mapping, identification of mutations, and identification of genetic disorders as well as other traits and to assess

PT biodiversity.

PS Claim 1; SEQ ID NO 874; 103pp; English

XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, CC and in recombinant production of (II). The polynucleotides are also used CC in diagnostics as expressed sequence tags for identifying expressed CC genes. (I) is useful in gene therapy techniques to restore normal CC activity of (II) or to treat disease states involving (II). (II) is CC useful for generating antibodies against it, detecting or quantitating a CC polypeptide in tissue, as molecular weight markers and as a food CC supplement. (II) and its binding partners are useful in medical imaging CC of sites expressing (II). (I) and (II) are useful for treating disorders CC involving aberrant protein expression or biological activity. The CC polypeptide and polynucleotide sequences have applications in CC diagnostics, forensics, gene mapping, identification of mutations CC responsible for genetic disorders or other traits to assess biodiversity CC and to produce other types of data and products dependent on DNA and CC amino acid sequences. A6664197-A6894554 represent novel human diagnostic CC coding sequences of the invention. Note: The sequence data for this CC patent did not appear in the printed specification, but was obtained in CC electronic format directly from WIPO at CC [http://wipo.int/pub/published\\_pat\\_sequences](http://wipo.int/pub/published_pat_sequences)

SQ Sequence 3259 BP; 944 A; 695 C; 674 G; 946 T; 0 U; 0 Other;

Query Match	97.6%	Score 2046	DB 5	Length 3259
-------------	-------	------------	------	-------------

Matches 2096; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGAGGAGTTTCTGTTACTATATGCTACACAGCAGGACAGGCAAGGCCATCGCAGAA 60

Db 80 ATGAGGAGTTCTGTTACTATATGCTACACAGCAGGACAGCAAGGCCATCGCAGAA 13

61 GAATATGTCAGCAAGCTGTGTACATGATTTCTGCAGATCTTCACCTGATTAGTGAA 12

Db 140 GAAATGCTGAGCAAGCTGTGTACATGGATTCTCTGCAGATCTTCACCTGATTAGTGAA 19

121 TCCGATAAGTATGACCTAAAAACCGAAACAGCTCCTCTTGTGTGTGTTCTACCACG 18

Db 200 TCCGATAAGTATGACCTAAAAACCGAACAGCTCCTCTTGTTGTTGTTCTACACG 250

181 GGCACCGAGACCCGACACAGCCCGCAAGTTGTTAAGGAATACAGAACCAACA 24

Db 260 GGCACCGAGACCCGACACAGCCCGCAAGTTGTTAAGGAATACAGAACCAACA 31

241 CTGCCGTTGATTTCTTGTCTACCTGCGGTATGGTTACTGGTCTCGGTATCAGAA 30

Db 320 CTGCCGGTGAATTCCTTGTCTACCTGCGGTATGGTTACTGGGCTCGGTGATTCAGAA 370

301 TACACCTACTTTTGCATGGGGGAGATATTGATAAAGACTTCAAGAGCTTGAGCC 36

Db 380 TACACCTACTTTTGCATGGGGGAGATATTGATAAAGACTTCAAGAGCTTGAGCC 43

361 CGGCA TTCTATGACACTGGACATGCAGATGACTGTGTAGGTTTAGAACTTGTGGTTGAG 42

Db 440 CGGCA TTTCTATGACACTGGACATGCAGATGACTGTGTAGGTTAGAACTTGTGTTGAG 495

421 CCGTGATTGCTGACTCTGGCCAGCCCTCAGAAAGCATTTTAGGTCAAGCAGAGGACAA 48

Db 500 CCGTGATTGCTGACTCTGGCCAGCCCTCAGAAGCATTTTAGTCAAGCAGAGGACAA 500

481 GAGGAGATAAGTGGCGCACTCCCGGTGGCATCACCTGCATCCTTGAGGACAGACCTTGTG 54

Db 560 GAGGAGATAAGTGGCGCACTCCGGTGGCATCACCTGCATCCTTGAGGACAGACCCTTGTG 61

541 AAGTCAGAGCTGCTACACATTGAATCTCAAGTCGAGCTTCTGAGATTGCATGATTCAGGA 60

Db 620 AAGTCAGAGCTGCTACACATTGAATCTCAAGTCGAGCTTCTGAGATTGATGATTCAGGA 67

601 AGAAGGATTCTGAGGTTTGAAGCAAAATGCAGTGAACAGCAACCAATCCAATGTTGTA 66

Db 680 AGAAGGATTCTGAGGTTTGAAGCAAAATGCAGTGAACAGCAACCAATCCAATGTTGTA 73



QY 661 ATTGAAGCTTTGAGTCTCACTTACCCTTCGGATCCCACTCTCAACAAGCTCTCTG 720  
DB 740 ATTGAAGCTTTGAGTCTCACTTACCCTTCGGATCCCACTCTCTCAACAAGCTCTCTG 729  
QY 721 AATATCTGTTTAACTCCCAAGAAATTTTACAGGTACATCGACAGAGTCTCTTGGCAG 780  
DB 800 AATATCTGTTTAACTCCCAAGAAATTTTACAGGTACATCGACAGAGTCTCTTGGCAG 859  
QY 781 GAGAAAGCCAAAGTATCTGACTTCCAGAGATCCAGTTTTCAGAGTCCAAATTTCAAG 840  
DB 860 GAGAAAGCCAAAGTATCTGACTTCCAGAGATCCAGTTTTCAGAGTCCAAATTTCAAG 919  
QY 841 GAGTTTCACTTACAGAAATGATGCAATTAACCACTGCTGCTGATGAAATTTGACAT 900  
DB 920 GAGTTTCACTTACAGAAATGATGCAATTAACCACTGCTGCTGATGAAATTTGACAT 979  
QY 901 TCAAAATCAGACTTTTCTATCAGCTGAGATGCTTACAGCTGATCTGCTTCAAGT 960  
DB 980 TCAAAATCAGACTTTTCTATCAGCTGAGATGCTTACAGCTGATCTGCTTCAAGT 1039  
QY 961 GATTTCTGAGTCAAAAGCTTACTTCAAAAGCTGAGCTTGAAGTAAAGAGACATG 1020  
DB 1040 GATTTCTGAGTCAAAAGCTTACTTCAAAAGCTGAGCTTGAAGTAAAGAGACATG 1099  
QY 1021 GTCTTTTGAATAAAGGAGACACAAAGAAAGAGCTACTTACCAGCATATA 1080  
DB 1100 GTCTTTTGAATAAAGGAGACACAAAGAAAGAGCTACTTACCAGCATATA 1159  
QY 1081 CTTGCGGAGATGTTCTCTCAAGTCAATTTTACCTGCTGTAAGTCCGAGCAATTCCT 1140  
DB 1160 CTTGCGGAGATGTTCTCTCAAGTCAATTTTACCTGCTGTAAGTCCGAGCAATTCCT 1219  
QY 1141 AAAAAGCATTTTTCGAGCCCTTGTGACTATAACAAGTACAGTGTGAAAAGCCGAG 1200  
DB 1220 AAAAAGCATTTTTCGAGCCCTTGTGACTATAACAAGTACAGTGTGAAAAGCCGAG 1279  
QY 1201 CTACAGAGCTGTGCACTTAAACAAGGGGCGAGCCGATTAATAGCCCTTTTGAACAAGTCC 1260  
DB 1280 CTACAGAGCTGTGCACTTAAACAAGGGGCGAGCCGATTAATAGCCCTTTTGAACAAGTCC 1339  
QY 1261 TGTGCTGCTGTTGATCTCTCTCTGCTTCCCTTCTTGCCAGCCAGCACTCAATCTC 1320  
DB 1340 TGTGCTGCTGTTGATCTCTCTCTGCTTCCCTTCTTGCCAGCCAGCACTCAATCTC 1399  
QY 1321 CTGCTGCAACATCTTCTTAACTTCAACCCAGACCATATTTGTGTGAGCTCAAGTTTA 1380  
DB 1400 CTGCTGCAACATCTTCTTAACTTCAACCCAGACCATATTTGTGTGAGCTCAAGTTTA 1459  
QY 1381 TTTTCAACCAAGAAAGCTTCCATTTTGTCTTCAACAATTTGTGAATTTCTGTCTACCTG 1440  
DB 1460 TTTTCAACCAAGAAAGCTTCCATTTTGTCTTCAACAATTTGTGAATTTCTGTCTACCTG 1519  
QY 1441 ACAGAGGTTCTGCGGAAGGAGTATGTACAGGCTGCGCTTGTGTTGCTTCAAGT 1500  
DB 1520 ACAGAGGTTCTGCGGAAGGAGTATGTACAGGCTGCGCTTGTGTTGCTTCAAGT 1579  
QY 1501 CTTCAGCCAAACATATGATGCAATCCATGAGAGACGCGGAAAGCCCTGCTCTTAAGATA 1560  
DB 1580 CTTCAGCCAAACATATGATGCAATCCATGAGAGACGCGGAAAGCCCTGCTCTTAAGATA 1639  
QY 1561 TCCATCTCTCTCGAACAACAATTTTTCATCTTACAGATGACCCCTCAATCCCATC 1620  
DB 1640 TCCATCTCTCTCGAACAACAATTTTTCATCTTACAGATGACCCCTCAATCCCATC 1699  
QY 1621 AATAATGAGGTTCCAGAACCGGATAGCCCGTTTATTTGGGTTCTTAAACATAGAG 1680  
DB 1700 AATAATGAGGTTCCAGAACCGGATAGCCCGTTTATTTGGGTTCTTAAACATAGAG 1759  
QY 1681 AAATCTCAAGAAACAACCCAGATGAGAAATTTTGGAGCAATGTGTTTGTGCTGC 1740  
DB 1760 AAATCTCAAGAAACAACCCAGATGAGAAATTTTGGAGCAATGTGTTTGTGCTGC 1819  
QY 1741 AGGCAATAGGATAGGATATCTATTCAAGAAAGAGCTCAGACATTTCTTAAGCATGG 1800

DB 1820 AGGCAATAGGATAGGATATCTATTCAAGAAAGAGCTCAGACATTTCTTAAGCATGG 1879  
QY 1801 ACTTAATCACTTAAAGGTTTCTCTCAAGAGATGCTCTGTTGGGAGAGAGAGCC 1860  
DB 1880 ACTTAATCACTTAAAGGTTTCTCTCAAGAGATGCTCTGTTGGGAGAGAGAGCC 1939  
QY 1861 CCAGCAAGTATGTACAGACCAATCAGCTTCAATGCGCAGAGGTGCGAGATCTC 1920  
DB 1940 CCAGCAAGTATGTACAGACCAATCAGCTTCAATGCGCAGAGGTGCGAGATCTC 1999  
QY 1921 CTCAGAGAGAGGCAATTTATGTTGTGAGATGCAAGAAATATGGCAAGATGT 1980  
DB 2000 CTCAGAGAGAGGCAATTTATGTTGTGAGATGCAAGAAATATGGCAAGATGT 2059  
QY 1981 CATGATGCTTGTGCAATTAATTAACCAAGAGTGTGAGTTGAAAACTAGAACATG 2040  
DB 2060 CATGATGCTTGTGCAATTAATTAACCAAGAGTGTGAGTTGAAAACTAGAACATG 2119  
QY 2041 AAAACCTGCGCATTTTAAAGAAAGAAACGCTACTTCAAGATATTTGTCTATA 2097  
DB 2120 AAAACCTGCGCATTTTAAAGAAAGAAACGCTACTTCAAGATATTTGTCTATA 2176

## RESULT 3

AAC91226 standard; DNA; 3259 BP.

AAC91226;

20-MAR-2001 (first entry)

Human schizophrenia related gene SEQ ID NO: 23.

Human: schizophrenia; developmental disorder; spina bifida cystica;

Kw Tourette's syndrome; bipolar illness; autism; conduct disorder;

Kw attention deficit disorder; obsessive compulsive disorder;

Kw chronic multiple tic syndrome; learning disorder; polymorphism; ds.

OS Homo sapiens.

PN MO200071754-A1.

30-NOV-2000.

24-MAY-2000; 2000MO-US014354.

25-MAY-1999; 99US-00318448.

(UTNE-) UNIV NEW JERSEY MEDICINE &amp; DENTISTRY.

Johnson WG, Stenroos ES;

WPI; 2001-025174/03.

Diagnosing a developmental disorder, e.g. schizophrenia, by forming

PT datasets (DS) of genetic (e.g. genotypes of folate metabolism alleles)

PT and environmental variables affecting an individual and then comparing

PT these DS with reference DS.

PS Disclosure; Page 142-143; 156pp; English.

CC The present invention provides a novel method of estimating the

CC susceptibility of an individual to a developmental disorder using genetic

CC prevention and treatment of disorders such as schizophrenia, spina bifida

CC cystica, Tourette's syndrome, bipolar illness, autism, conduct disorders,

CC attention deficit disorder, obsessive compulsive disorder, chronic

CC multiple tic syndrome and learning disorders such as dyslexia

Query Match

97.6%; Score 2046; DB 5; Length 3259;

Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2096; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGAGGTTTCTGTACTATATGCTACACAGCAGGAGCAGCAAAAGCCATCGCAGAA 60  
DB |||||  
QY 80 ATGAGAGGTTTCTGTACTATATGCTACACAGCAGGAGCAGCAAAAGCCATCGCAGAA 139  
DB |||||  
QY 61 GAAATATGAGCAGCTGTGTACATGATTTCTGAGATCTTCACTGTATTAAGTAA 120  
DB |||||  
QY 140 GAAATGAGCAGCTGTGTACATGATTTCTGAGATCTTCACTGTATTAAGTAA 199  
DB |||||  
QY 121 TCCGATTAAGTATGACCTTAAACCCGAAAGAGCTCTCTGTGTGTGTTTCTACAG 180  
DB |||||  
QY 200 TCGATTAAGTATGACCTTAAACCCGAAAGAGCTCTCTGTGTGTGTTTCTACAG 259  
DB |||||  
QY 181 GGCACCGGAGACCCGACACAGCCGAGTTGTTAAGAAATACAGAACCAACA 240  
DB |||||  
QY 260 GGCACCGGAGACCCGACACAGCCGAGTTGTTAAGAAATACAGAACCAACA 319  
DB |||||  
QY 241 CTGCGGTTGATTTCTGTCTACCTGCGGTATGAGTTAAGTATGAGTTTGA 300  
DB |||||  
QY 320 CTGCGGTTGATTTCTGTCTACCTGCGGTATGAGTTAAGTATGAGTTTGA 379  
DB |||||  
QY 301 TACACCTACTTTTGAATGAGGAGGAGATTAATGATTAACGACTTCAAGAGCTTGAAGC 360  
DB |||||  
QY 380 TACACCTACTTTTGAATGAGGAGGAGATTAATGATTAACGACTTCAAGAGCTTGAAGC 439  
DB |||||  
QY 361 CGGACTTTCTATGACATGAGACATGATGATGATGATTTAGAACTTGTGTTGAG 420  
DB |||||  
QY 440 CGGACTTTCTATGACATGAGACATGATGATGATGATTTAGAACTTGTGTTGAG 499  
DB |||||  
QY 421 CCGTGAATGCTGAGATCTGTGCGGACCTTCAAGAACTTTTGAAGCAGAGAGCA 480  
DB |||||  
QY 500 CCGTGAATGCTGAGATCTGTGCGGACCTTCAAGAACTTTTGAAGCAGAGAGCA 559  
DB |||||  
QY 481 GAGAGATATGAGGCGCAGCTCCCGGTGAGATCACTGACCTTCTGAGAGCAGAGCTTGTG 540  
DB |||||  
QY 560 GAGAGATATGAGGCGCAGCTCCCGGTGAGATCACTGACCTTCTGAGAGCAGAGCTTGTG 619  
DB |||||  
QY 541 AAGTCAGAGCTGTACACATTTGATATCTCAAGTCGATCTTGAATTCATGATTCAGAA 600  
DB |||||  
QY 620 AAGTCAGAGCTGTACACATTTGATATCTCAAGTCGATCTTGAATTCATGATTCAGAA 679  
DB |||||  
QY 601 AGAAAGATTTCTGAGTTTGAAGCAAAATGCAAGTAAACAGCAACCAATCCAAATGTTTGA 660  
DB |||||  
QY 680 AGAAAGATTTCTGAGTTTGAAGCAAAATGCAAGTAAACAGCAACCAATCCAAATGTTTGA 739  
DB |||||  
QY 661 ATTTGAGATTTGAGTCTCACTTACCCGTTGAGTACCCCACTCTCAAGAGCTTCTG 720  
DB |||||  
QY 740 ATTTGAGATTTGAGTCTCACTTACCCGTTGAGTACCCCACTCTCAAGAGCTTCTG 799  
DB |||||  
QY 721 AATATTCCTGTGTTTACCCCAAGATATTTTACAGTATCATCTGAGAGAGTCTTGTGCGAG 780  
DB |||||  
QY 800 AATATTCCTGTGTTTACCCCAAGATATTTTACAGTATCATCTGAGAGAGTCTTGTGCGAG 859  
DB |||||  
QY 781 GAGGAAAGCAAGTATCTGTGATTCAGCAGATCCAGTTTCAAGTCCCAATTTCAAG 840  
DB |||||  
QY 860 GAGGAAAGCAAGTATCTGTGATTCAGCAGATCCAGTTTCAAGTCCCAATTTCAAG 919  
DB |||||  
QY 841 GAGGTTCAACTTCTCAAGATGATGCAATTAACCACTCTGCTGAGATTTGAGCATT 900  
DB |||||  
QY 920 GAGGTTCAACTTCTCAAGATGATGCAATTAACCACTCTGCTGAGATTTGAGCATT 979  
DB |||||  
QY 901 TCAATATACAGCTTTTCTATCAGCCTGAGATGCTTCAAGCCTGATCTGCTTAA 960  
DB |||||  
QY 980 TCAATATACAGCTTTTCTATCAGCCTGAGATGCTTCAAGCCTGATCTGCTTAA 1039  
DB |||||  
QY 961 GATTTGAGGTTACAAAGCTTCTCAAGATGCTGAGTTTGAAGTAAAGAGAGCAGCTGC 1020  
DB |||||  
QY 1040 GATTTGAGGTTACAAAGCTTCTCAAGATGCTGAGTTTGAAGTAAAGAGAGCAGCTGC 1099  
DB |||||  
QY 1021 GTCTTTTGAATTAAGGAGACACAAAGAAAGAGAGCTTACCCAGCATATA 1080  
DB |||||

DB 1100 GTCTTTTGAATTAAGGAGACACAAAGAAAGAGAGCTTACCCAGCATATA 1159  
QY |||||  
DB 1081 CCTGGGGAATGTTCTTCAGTTCAATTTTACCTGAGTCTGAATCCAGCAATTCCT 1140  
QY |||||  
DB 1160 CCGGAGGAGTGTCTTCAGTTCAATTTTACCTGAGTCTGAATCCAGCAATTCCT 1219  
QY |||||  
DB 1141 AAAAAGCATTTTTTCGAGCCCTTGTGAGCTTATACAGTGAAGTGTGAAAAGCAGAG 1200  
QY |||||  
DB 1220 AAAAAGCATTTTTTCGAGCCCTTGTGAGCTTATACAGTGAAGTGTGAAAAGCAGAG 1279  
QY |||||  
DB 1201 CTACAGAGCTGTGAGTAAACAGAGGAGCCGATTAATAGCCGTTTGTAGAGATGCC 1260  
QY |||||  
DB 1280 CTACAGAGCTGTGAGTAAACAGAGGAGCCGATTAATAGCCGTTTGTAGAGATGCC 1339  
QY |||||  
DB 1261 TGTGCTGTGTGAGTATCTCTCTGCTTCCCTTCTGTGAGCAGCAGCAGCTGAGTCTC 1320  
QY |||||  
DB 1340 TGTGCTGTGTGAGTATCTCTCTGCTTCCCTTCTGTGAGCAGCAGCAGCTGAGTCTC 1399  
QY |||||  
DB 1321 CTGCTGAAATCTTCTTAACTTCAACCCAGACATATTCGTGACAGCTCAAGTTTA 1380  
QY |||||  
DB 1400 CTGCTGAAATCTTCTTAACTTCAACCCAGACATATTCGTGACAGCTCAAGTTTA 1459  
QY |||||  
DB 1381 TTTCAACCGAGAAAGCTCCATTTTGTCTTCAATTTGTGAAATTTCTGTACTGACACA 1440  
QY |||||  
DB 1460 TTTCAACCGAGAAAGCTCCATTTTGTCTTCAATTTGTGAAATTTCTGTACTGACACA 1519  
QY |||||  
DB 1441 ACAGAGTTCTGCGAGAGGAGATATGATAGCTGCTGAGCTTGTGTGTTGCTTCAATT 1500  
QY |||||  
DB 1520 ACAGAGTTCTGCGAGAGGAGATATGATAGCTGCTGAGCTTGTGTGTTGCTTCAATT 1579  
QY |||||  
DB 1501 CTTGAGCCAAATATCATGATCCCATGAAAGACAGGAGGAAAGCCCTGAGCTTAAAGATA 1560  
QY |||||  
DB 1580 CTTGAGCCAAATATCATGATCCCATGAAAGACAGGAGGAAAGCCCTGAGCTTAAAGATA 1639  
QY |||||  
DB 1561 TCCATCTCTCTGAGCAACAAATTTCTTCACTTACAGATGAGCCCTCAATCCCATC 1620  
QY |||||  
DB 1640 TCCATCTCTCTGAGCAACAAATTTCTTCACTTACAGATGAGCCCTCAATCCCATC 1699  
QY |||||  
DB 1621 ATTAATGTTGGTTCAGAGAACCGGATAGCCCGTTTATTTGGGTTCCTTAAACAATAGAG 1680  
QY |||||  
DB 1700 ATTAATGTTGGTTCAGAGAACCGGATAGCCCGTTTATTTGGGTTCCTTAAACAATAGAG 1759  
QY |||||  
DB 1681 AAACTCAAGAAACAACACCAAGTGAATTTTGTGAGCATGTGTTGTTTGTGCTGC 1740  
QY |||||  
DB 1760 AAACTCAAGAAACAACACCAAGTGAATTTTGTGAGCATGTGTTGTTTGTGCTGC 1819  
QY |||||  
DB 1741 AAGCATTAAGATTAAGGATTAATCTATTCAGAAAAGAGCTCAGACATTTCTTAAGCATGAG 1800  
QY |||||  
DB 1820 AAGCATTAAGATTAAGGATTAATCTATTCAGAAAAGAGCTCAGACATTTCTTAAGCATGAG 1879  
QY |||||  
DB 1801 ATCTTAATCTATTAAGGTTTCTTCTCAAGAGATGCTCTGTGTGAGAGAGAGAGCC 1860  
QY |||||  
DB 1880 ATCTTAATCTATTAAGGTTTCTTCTCAAGAGATGCTCTGTGTGAGAGAGAGAGCC 1939  
QY |||||  
DB 1861 CCAGCAAGATTAATTAAGCAACACATCAGCTTCAATGAGCAGAGGTGAGAGATCTCTC 1920  
QY |||||  
DB 1940 CCAGCAAGATTAATTAAGCAACACATCAGCTTCAATGAGCAGAGGTGAGAGATCTCTC 1999  
QY |||||  
DB 2000 CTCCAGAGAAAGGCGCATTTTATGTGTGTGAGATGCAAAAGATATGCGCAAGATGTA 2059  
QY |||||  
DB 1981 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGGTTGAGTTGAAAATCTAAGAGCAATG 2040  
QY |||||  
DB 2060 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGGTTGAGTTGAAAATCTAAGAGCAATG 2119  
QY |||||  
DB 2041 AAAAACCCTGCGCACTTTAAAGAGAAAACGCTACAGGATTTTGTGCTATA 2097  
QY |||||  
DB 2120 AAAAACCCTGCGCACTTTAAAGAGAAAACGCTACAGGATTTTGTGCTATA 2176

RESULT 4  
ADM43206

ID ADM43206 standard; cDNA; 3259 BP.  
XX  
AC ADM43206;  
XX  
XX  
DT 03-JUN-2004 (first entry)  
XX  
DE Human full length cDNA encoding methionine synthase reductase.  
XX  
XX Human; ss; gene: Methionine synthase reductase polypeptide; HmTRR;  
KM cancer; cardiovascular disease; neural tube defect;  
KM hypethomocysteinaemia; chromosome 5p15.2-p15.3; SNP;  
XX single nucleotide polymorphism.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 80..2176  
FT /\*tag= a  
FT /product= "HmTRR"  
FT /replace(145,A)  
FT /\*tag= b  
FT /strand.name= "Single\_nucleotide\_polymorphism"  
FT /replace(189,A)  
FT variation  
FT /tag= c  
FT /standard\_name= "Single\_nucleotide\_polymorphism"  
XX  
PN US2003082676-A1.  
XX  
PD 01-MAY-2003.  
XX  
XX 10-AUG-1999; 99US-00371347.  
XX  
XX 15-JAN-1998; 98US-0071622P.  
PR 16-JAN-1999; 99US-00232028.  
XX  
XX (GRAY/) GRAYEL R A.  
PA (ROZE/) ROZEN R.  
PA (LECL/) LECLERC D.  
PA (WILS/) WILSON A.  
PA (ROSE/) ROSENBLATT D.  
XX  
PI Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;  
XX  
XX WPI; 2003-57610/54.  
DR P-PSDB; ADM43207.  
XX  
PT New substantially pure nucleic acid encoding a mammalian methionine  
PT synthase reductase polypeptide, useful for diagnosing, preventing or  
PT treating conditions associated with altered methionine synthase activity,  
PT e.g. cancer.  
XX  
XX Example 2; SEQ ID NO 24; 26pp; English.  
XX  
XX The invention relates to a substantially pure nucleic acid that encodes a  
XX mammalian methionine synthase reductase polypeptide, HmTRR, or that  
XX hybridizes at high stringency to a nucleic acid appearing as ADM43208 or  
XX ADM43209. Also included are a non-human animal where one or both genetic  
XX alleles encoding the methionine synthase reductase polypeptide are  
XX mutated, an antibody that specifically binds the above methionine  
XX synthase reductase polypeptide, a method of detecting the presence of the  
XX above polypeptide, a method for detecting sequence variants for  
XX methionine synthase reductase in a mammal, methods of treating or  
XX preventing cancer (or cardiovascular disease or neural tube defects) in a  
XX subject, methods of screening for a compound that modulates methionine  
XX synthase reductase biological activity and a method for detecting an  
XX increased risk of developing a neural tube defect in a mammalian embryo  
XX or foetus. The nucleic acid is useful in diagnosing, preventing or  
XX treating conditions associated with altered methionine synthase activity,  
XX such as cancer, cardiovascular disease or neural tube defects, or in  
XX screening for a compound that modulates methionine synthase reductase  
XX biological activity. Naturally occurring variants of the polypeptide are  
XX also associated with hypethomocysteinaemia. The gene for HmTRR is  
XX located on chromosome 5p15.2-p15.3. The present sequence is full length

CC sequence of the wild-type human hwmTRR cDNA.  
XX  
SQ Sequence 3259 BP; 944 A; 695 C; 674 G; 946 T; 0 U; 0 Other;  
XX  
Query Match 97.6%; Score 2046; DB 11; Length 3259;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2096; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ATGAGAGAGTTTCTGTATCATATGCTATGACAGACAGGACGCAAGCCATCGCAGAA 60  
DB 80 ATGAGAGAGTTTCTGTATCATATGCTATGACAGACAGGACGCAAGCCATCGCAGAA 139  
QY 61 GAAATATGTGACCAAGCTGTGTACATGGAATTTCTGCAATCTTCATGTATTAAGTAA 120  
DB 140 GAAATGTGACCAAGCTGTGTACATGGAATTTCTGCAATCTTCATGTATTAAGTAA 199  
QY 121 TCGATTAAGTATGACCTATAAAACCGAAACAGCTCCCTGTGTGTGTGTCTTACACAG 180  
DB 200 TCGATTAAGTATGACCTATAAAACCGAAACAGCTCCCTGTGTGTGTGTCTTACACAG 259  
QY 181 GGCACCGAGACCCACCCGACACAGCCCGCAAGTTGTTAAGGAATTAAGAACCAACA 240  
DB 260 GGCACCGAGACCCACCCGACACAGCCCGCAAGTTGTTAAGGAATTAAGAACCAACA 319  
QY 241 CTGCCGTTGATTTCTTGTCTCACTGCGGTATGAGTTA CTGGGTCTCGGTATTCAGAA 300  
DB 320 CTGCCGTTGATTTCTTGTCTCACTGCGGTATGAGTTA CTGGGTCTCGGTATTCAGAA 379  
QY 301 TACACCTACTTTTGGCAATGGGGGAGATTAATGATTAACGCTTAAGAGCTTGAGCC 360  
DB 380 TACACCTACTTTTGGCAATGGGGGAGATTAATGATTAACGCTTAAGAGCTTGAGCC 439  
QY 361 CGCATTTCTATGACACTGGAACATGACATGACTGTGTAGTTTGAACCTTGTGTGAG 420  
DB 440 CGCATTTCTATGACACTGGAACATGACATGACTGTGTAGTTTGAACCTTGTGTGAG 499  
QY 421 CGGTGATTTGCTGACTCTGGCCAGCCCTCAGAAAGCATTTTATGTCAGACAGACAA 480  
DB 500 CGGTGATTTGCTGACTCTGGCCAGCCCTCAGAAAGCATTTTATGTCAGACAGACAA 559  
QY 481 GAGGAGATTAAGGGGCGCACTCCCGGTGGCATCACTGCAATCCTTGAAGGACAGACTTGG 540  
DB 560 GAGGAGATTAAGGGGCGCACTCCCGGTGGCATCACTGCAATCCTTGAAGGAGACTTGG 619  
QY 541 AAGTCAGAGCTGCTACACATTTGAATCTCAAGTGTGAGCTTCTGAGATTCTGATTCAGGA 600  
DB 620 AAGTCAGAGCTGCTACACATTTGAATCTCAAGTGTGAGCTTCTGAGATTCTGATTCAGGA 679  
QY 601 AGAAGGATTTGAGGTTTGAAGCAAAATGCAAGTAACAGCAACCAATCCAAATGTTGTA 660  
DB 680 AGAAGGATTTGAGGTTTGAAGCAAAATGCAAGTAACAGCAACCAATCCAAATGTTGTA 739  
QY 661 ATTGAAGACTTTGAGTCCTACCTACCCGTGGTACCCCACTTCACAAAGCTCTCG 720  
DB 740 ATTGAAGACTTTGAGTCCTACCTACCCGTGGTACCCCACTTCACAAAGCTCTCG 799  
QY 721 AATATTCTGTGTTTAAACCCCAAGAAATTTTACAGGTACATCTGACAGAGTCTCTTGCCAG 780  
DB 800 AATATTCTGTGTTTAAACCCCAAGAAATTTTACAGGTACATCTGACAGAGTCTCTTGCCAG 859  
QY 781 GAGGAAGCCAGATATCTGTGACTTCAGACAGATTCAGATTTTCAAGTCCAAATTCAGAG 840  
DB 860 GAGGAAGCCAGATATCTGTGACTTCAGACAGATTCAGATTTTCAAGTCCAAATTCAGAG 919  
QY 841 GCGATTCACCTTAATCTGCAATGATGCAATTAACCACTGTGTGATTAATGAGCATT 900  
DB 920 GCGATTCACCTTAATCTGCAATGATGCAATTAACCACTGTGTGATTAATGAGCATT 979  
QY 901 TCAATATACAGACTTTTCTATACGCTTGAGATGCTTCAAGCTGTATCTGCCATACAGT 960  
DB 980 TCAATATACAGACTTTTCTATACGCTTGAGATGCTTCAAGCTGTATCTGCCATACAGT 1039  
QY 961 GATTCTGAGTACAAAGCTTACCAAGACTGCAAGCTGGAAGATTAAGAGAGCACTGC 1020

Db 1040 GATTCTGAGGTCAAGACCTACTCAAGAGCTGCACTTGAAGATTAAGAGAGCACTGC 1099  
 QY 1021 GTCTCTTTTAAAAATTAAGGACAGACAAAGAAAGAGAGCTTACTTACCCAGCATATA 1080  
 Db 1100 GTCTCTTTTAAAAATTAAGGACAGACAAAGAAAGAGAGCTTACTTACCCAGCATATA 1159  
 QY 1081 CTGCGGAGATGTCTCTCAAGTTCATTTTACTGTGTCTTGAATTCGAGCATTTCT 1140  
 Db 1160 CTGCGGAGATGTCTCTCAAGTTCATTTTACTGTGTCTTGAATTCGAGCATTTCT 1219  
 QY 1141 AAAAAGGCAATTTTGGAGCCCTTGTGACTATACAGTGAAGAGCTGAAAAAGCCAG 1200  
 Db 1220 AAAAAGGCAATTTTGGAGCCCTTGTGACTATACAGTGAAGAGCTGAAAAAGCCAG 1279  
 QY 1201 CTACAGAGCTGTGCAAGTAAACAAAGGAGGAGCCGATTAATAGCCGCTTTGTAGAGATGCC 1260  
 Db 1280 CTACAGAGCTGTGCAAGTAAACAAAGGAGGAGCCGATTAATAGCCGCTTTGTAGAGATGCC 1339  
 QY 1261 TGGAGCTGTGTGTGATGCTCCCTCGGCTTCCCTTTCGAGGACACCACTCACTCTC 1320  
 Db 1340 TGGAGCTGTGTGTGATGCTCCCTCGGCTTCCCTTTCGAGGACACCACTCACTCTC 1399  
 QY 1321 CTGCTCGAATCTTCTCTAACTTCAACCCAGACATATTCGTGCAAGCTCAAGTTTA 1380  
 Db 1400 CTGCTCGAATCTTCTCTAACTTCAACCCAGACATATTCGTGCAAGCTCAAGTTTA 1459  
 QY 1381 TTTCACCCAGAAAGCTCAATTTTGTCTTCAACATTTGTGGAATTTCTGTCTACTGCCACA 1440  
 Db 1460 TTTCACCCAGAAAGCTCAATTTTGTCTTCAACATTTGTGGAATTTCTGTCTACTGCCACA 1519  
 QY 1441 ACAGAGGTTCTGCGGAAAGGAGATATGATACAGCTGAGCTGCTGTGTGTGCTTCAAGTT 1500  
 Db 1520 ACAGAGGTTCTGCGGAAAGGAGATATGATACAGCTGAGCTGCTGTGTGTGCTTCAAGTT 1579  
 QY 1501 CTTCAGCCAAACATACATGATCTCCATGAGAAGAGGAGAAAGCCCTGCTCCTTAAGATA 1560  
 Db 1580 CTTCAGCCAAACATACATGATCTCCATGAGAAGAGGAGAAAGCCCTGCTCCTTAAGATA 1639  
 QY 1561 TCGATCTCTCTCTGAAACAACAATTTCTTTCACATTAACAGATGACCCCTCAATCCCAATC 1620  
 Db 1640 TCGATCTCTCTCTGAAACAACAATTTCTTTCACATTAACAGATGACCCCTCAATCCCAATC 1699  
 QY 1621 ATTAATGTGGTCCAGAGAACCCGATAGCCCGTTATGTGGTCTTACAAACATAGAGAG 1680  
 Db 1700 ATTAATGTGGTCCAGAGAACCCGATAGCCCGTTATGTGGTCTTACAAACATAGAGAG 1759  
 QY 1681 AAATCTCAAGAACAAACACCCAGATGAAATTTTGAAGCAATGTGTTGTTTTTGGCTGC 1740  
 Db 1760 AAATCTCAAGAACAAACACCCAGATGAAATTTTGAAGCAATGTGTTGTTTTTGGCTGC 1819  
 QY 1741 AGGCATTAAGATAGGATTAATCTATTCAAGAAAAGACTCAGACATTTCTTAAAGATGG 1800  
 Db 1820 AGGCATTAAGATAGGATTAATCTATTCAAGAAAAGACTCAGACATTTCTTAAAGATGG 1879  
 QY 1801 ATCTTAACATCTTAAGGTTTCTCTCAAGAGATGCTCTGTGTGGAGAGGAGAACCC 1860  
 Db 1880 ATCTTAACATCTTAAGGTTTCTCTCTCAAGAGATGCTCTGTGTGGAGAGGAGAACCC 1939  
 QY 1861 CCAGCAAAATATGTACAAGCAACATCAGCTTCATGSCCAGAGAGGTGGAGAAATCTCTC 1920  
 Db 1940 CCAGCAAAATATGTACAAGCAACATCAGCTTCATGSCCAGAGAGGTGGAGAAATCTCTC 1999  
 QY 1921 CTCACAGAGAACGCGCATATTTATGTGTGTGAGATGCAAGAAATATGCGCAAGATGTA 1980  
 Db 2000 CTCACAGAGAACGCGCATATTTATGTGTGTGAGATGCAAGAAATATGCGCAAGATGTA 2059  
 QY 1981 CATTGATGCTTGTGCAAAATTAATACAAAGAGTTGAGTTGAAAAAATCAGAAAGCAATG 2040  
 Db 2060 CATTGATGCTTGTGCAAAATTAATACAAAGAGTTGAGTTGAAAAAATCAGAAAGCAATG 2119  
 QY 2041 AAAACCTGCGCACTTTAAAAAGAAAGAAACGCTACCTTCAGAGATTTTGTGATATA 2097

Db 2120 AAAACCTGCGCACTTTAAAAAGAAAGAAACGCTACCTTCAGAGATTTTGTGATATA 2176  
 RESULT 5  
 ADM43208  
 ID ADM43208 standard; cDNA; 2094 BP.  
 XX  
 AC ADM43208;  
 XX  
 DT 03-JUN-2004 (first entry)  
 XX  
 DE Human wild-type methionine synthase reductase CDS.  
 XX  
 KW Human; ss; Methionine synthase reductase polypeptide; HsmTRR; cancer;  
 KW cardiovascular disease; neural tube defect; hyperhomocysteinemia;  
 KW chromosome 5p15.2-p15.3; SNP; single nucleotide polymorphism.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..2094  
 FT /\*tag= a  
 FT /product= "hsmTRR"  
 FT /partial  
 FT /note= "No stop codon shown"  
 FT variation  
 FT /\*tag= b  
 FT /standard\_name= "Single\_nucleotide\_polymorphism"  
 FT replace(110,A)  
 FT /\*tag= c  
 FT /standard\_name= "Single\_nucleotide\_polymorphism"  
 XX  
 PN US2003082676-A1.  
 XX  
 PD 01-MAY-2003.  
 XX  
 PF 10-AUG-1999; 99US-00371347.  
 XX  
 PR 16-JAN-1998; 98US-0071622P.  
 PR 15-JAN-1999; 99US-00232028.  
 XX  
 PA (GRAV/) GRAVEL R A.  
 PA (ROZE/) ROZEN R.  
 PA (LECL/) LECLERC D.  
 PA (WILS/) WILSON A.  
 PA (ROSE/) ROSENBLATT D.  
 XX  
 PI Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;  
 XX  
 DR WPI: 2003-576610/54.  
 DR P-PSDB; ADM43207.  
 XX  
 PT New substantially pure nucleic acid encoding a mammalian methionine  
 PT synthase reductase polypeptide, useful for diagnosing, preventing or  
 PT treating conditions associated with altered methionine synthase activity,  
 PT e.g. cancer.  
 XX  
 PS Claim 3; SEQ ID NO 1; 26pp; English.  
 XX  
 CC The invention relates to a substantially pure nucleic acid that encodes a  
 CC mammalian methionine synthase reductase polypeptide, HsmTRR, or that  
 CC hybridizes at high stringency to a nucleic acid appearing as ADM43208 or  
 CC ADM43209. Also included are a non-human animal where one or both genetic  
 CC alleles encoding the methionine synthase reductase polypeptide are  
 CC mutated, an antibody that specifically binds the above methionine  
 CC synthase reductase polypeptide, a method of detecting the presence of the  
 CC above polypeptide, a method for detecting sequence variants for  
 CC methionine synthase reductase in a mammal, methods of treating or  
 CC preventing cancer (or cardiovascular disease or neural tube defects) in a  
 CC subject, methods of screening for a compound that modulates methionine  
 CC synthase reductase biological activity and a method for detecting an  
 CC increased risk of developing a neural tube defect in a mammalian embryo  
 CC or foetus. The nucleic acid is useful in diagnosing, preventing or



QY 1981 CATGATGCCCTTGTGCAATTAATAGCAAGAGTTGAGTTGAAAAGTACAGAGCAATG 2040  
DB 1981 CATGATGCCCTTGTGCAATTAATAGCAAGAGTTGAGTTGAAAAGTACAGAGCAATG 2040  
QY 2041 AAAACCTGCGCACTTTAAAGAGAGAAACCGTACCTTCAGATATTTTGTCA 2094  
DB 2041 AAAACCTGCGCACTTTAAAGAGAGAAACCGTACCTTCAGATATTTTGTCA 2094

RESULT 6  
ADQ87538  
ID ADQ87538 standard; cDNA; 3270 BP.  
AC ADQ87538;  
XX  
XX  
XX 07-OCT-2004 (first entry)  
DT  
XX Human tumour-associated antigenic target (TAT) cDNA sequence #4416.  
DE human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;  
XX cancer; cell proliferative disorder; gene; ss.  
XX Homo sapiens.  
OS  
XX MO2004060270-A2.  
PN  
XX 22-JUL-2004.  
PD  
XX 15-OCT-2003; 2003MO-US029126.  
PF  
XX 18-OCT-2002; 2002US-0418988P.  
PR  
XX (GETH) GENENTECH INC.  
PA (WUTD/) WU T D.  
XX (ZHOU/) ZHOU Y.  
PI Wu TD, Zhou Y;  
XX WPI; 2004-534300/51.  
DR  
XX New nucleic acid molecule and encoded polypeptide, for diagnosing,  
PT preventing or treating cell proliferative disorders such as cancer.  
PS Claim 1; SEQ ID NO 4416; 5504bp; English.

CC The present invention describes an isolated tumour-associated antigenic  
CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide  
CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of  
CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%  
CC sequence identity to (a)-(c); or (e) a sequence that hybridizes to (a)-  
CC (c). Also described: (1) an expression vector comprising the above  
CC nucleic acid; (2) a host cell comprising the above expression vector; (3)  
CC a process for producing a polypeptide; (4) an isolated polypeptide  
CC comprising: (a) an amino acid sequence encoded by any of the above  
CC nucleotide sequences; (b) an amino acid sequence encoded by the full-  
CC length coding region of the above nucleotide sequences; or (c) a sequence  
CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide  
CC comprising the above polypeptide fused to a heterologous polypeptide; (6)  
CC an isolated antibody that binds to the above polypeptide; (7) a process  
CC for producing the antibody; (8) an isolated oligopeptide that binds to  
CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)  
CC binding organic molecule that binds to the above polypeptide; (10) a  
CC composition of matter comprising the above (chimeric) polypeptide,  
CC antibody, oligopeptide or TAT binding organic molecule, in combination  
CC with a carrier; (11) an article of manufacture comprising a container and  
CC the composition of matter contained within the container; (12) methods of  
CC inhibiting the growth of a cell that expresses the above protein, where  
CC the growth of the cell is at least in part dependent upon a growth  
CC potentiating effect of the above protein; (13) a method of  
CC therapeutically treating a mammal having a cancerous tumour comprising  
CC cells that express the above protein; (14) a method of determining the  
CC presence of a protein in a sample suspected of containing the protein  
CC described above; (15) methods of diagnosing the presence of a tumour in a

CC mammal; (16) a method for treating or preventing a cell proliferative  
CC disorder associated with increased expression or activity of the above  
CC protein; and (17) a method of binding an antibody, oligopeptide or  
CC organic molecule to a cell that expresses the protein described above.  
CC The TAT sequences have cytostatic activities, and can be used in gene  
CC therapy. The composition and methods are useful for diagnosing,  
CC preventing or treating cancer. The composition is also used for preparing  
CC a medicament for the therapeutic treatment or diagnostic detection of a  
CC cell proliferative disorder or cancer. The present sequence represents a  
CC human TAT cDNA sequence from the present invention.  
XX  
XX Sequence 3270 BP; 934 A; 702 C; 680 G; 954 T; 0 U; 0 Other;  
SQ  
Query Match 95.1%; Score 1995; DB 13; Length 3270;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2095; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGAGGTTTCTGTATCTATATGCTACACAGAGGACAGGCAAGCCATCCGAGAA 60  
DB 112 ATGAGAGGTTTCTGTATCTATATGCTACACAGAGGACAGGCAAGCCATCCGAGAA 171  
QY 61 GAAATATGAGACCAAGCTGTGTACATGATTTTCTGAGATCTTCACTATATGTA 120  
DB 172 GAAATATGAGACCAAGCTGTGTACATGATTTTCTGAGATCTTCACTATATGTA 231  
QY 121 TCCGATTAAGTATGACCTTAAACCGAAACAGCTCTCTGTGTGTGTGTTCTACACG 180  
DB 232 TCCGATTAAGTATGACCTTAAACCGAAACAGCTCTCTGTGTGTGTGTTCTACACG 291  
QY 181 GGCACCGAGAGACCAACCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCA 240  
DB 292 GGCACCGAGAGACCAACCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCA 351  
QY 241 CTCGGGTTGATTTCTTTGCTCACTGCGGTATGSGTTACTGGGTCTGGTATTCAGAA 300  
DB 352 CTCGGGTTGATTTCTTTGCTCACTGCGGTATGSGTTACTGGGTCTGGTATTCAGAA 411  
QY 301 TACACCTACTTTTGAATGGGGGGAAGATTAATGATTAACGACTTCAAGAGCTGGAGCC 360  
DB 412 TACACCTACTTTTGAATGGGGGGAAGATTAATGATTAACGACTTCAAGAGCTGGAGCC 471  
QY 361 CGCATTTCTATGACACTGACATGACATGATGATGATGATGATGATGATGATGATG 420  
DB 472 CGCATTTCTATGACACTGACATGACATGATGATGATGATGATGATGATGATGATG 531  
QY 421 CCGTGATTTCTGACCTGCGCAGCCCTCAGAAAGCATTTTAAAGTCAAGAGAGCA 480  
DB 532 CCGTGATTTCTGACCTGCGCAGCCCTCAGAAAGCATTTTAAAGTCAAGAGAGCA 591  
QY 481 GAGGAGATAGGGGGCACTCCGGTGATCACTGATCTGTAAGAGCAGACCTGTG 540  
DB 592 GAGGAGATAGGGGGCACTCCGGTGATCACTGATCTGTAAGAGCAGACCTGTG 651  
QY 541 AAGTCAGAGCTCTACATTAATCTCAAGTGAAGCTTGAAGATTCAGATTCAGAGA 600  
DB 652 AAGTCAGAGCTCTACATTAATCTCAAGTGAAGCTTGAAGATTCAGATTCAGAGA 711  
QY 601 AAAAAAGATTCGAGGTTTGAAGCAAAATGCAAGTACAGCAACCAATTCATTTGTA 660  
DB 712 AAAAAAGATTCGAGGTTTGAAGCAAAATGCAAGTACAGCAACCAATTCATTTGTA 771  
QY 661 ATTGAAGATTTGAGTCCCACTTACCCGTTGGTGAACCCCACTTCAGAAAGCTCTG 720  
DB 772 ATTGAAGATTTGAGTCCCACTTACCCGTTGGTGAACCCCACTTCAGAAAGCTCTG 831  
QY 721 AATATTCCTGTTTACCCCAAGATATTTTACAGTATCTGACAGAGTCTTGGCCAG 780  
DB 832 AATATTCCTGTTTACCCCAAGATATTTTACAGTATCTGACAGAGTCTTGGCCAG 891  
QY 781 GAGGAAAGCAGATATCTGATCTTCAAGATTCAGATTTTCAAGTGCATTTCAAG 840  
DB 892 GAGGAAAGCAGATATCTGATCTTCAAGATTCAGATTTTCAAGTGCATTTCAAG 951



QY 841 GCAGTCACTTATCTACGAATGATGCAATTAACCACTCTGCTGTAGAAATTGGACATT 900  
DB 952 GCAGTCACTTATCTACGAATGATGCAATTAACCACTCTGCTGTAGAAATTGGACATT 1011  
QY 901 TCAAAATACAACTTTTCTATCAGCCCTGAGAGCTTCAAGCCGTGATCTGCTTAACT 960  
DB 1012 TCAAAATACAACTTTTCTATCAGCCCTGAGAGCTTCAAGCCGTGATCTGCTTAACT 1071  
QY 961 GATTCTGAGTACAAAGCTTCTCAAGAATGAGCTTGAAGATTAAGAGAGCACTGC 1020  
DB 1072 GATTCTGAGTACAAAGCTTCTCAAGAATGAGCTTGAAGATTAAGAGAGCACTGC 1131  
QY 1021 GTCTCTTGAATAATTAAGGACACAAAGAGAGAGCTTACCTCCAGCATATA 1080  
DB 1132 GTCTCTTGAATAATTAAGGACACAAAGAGAGAGCTTACCTCCAGCATATA 1191  
QY 1081 CCTGCGGAGATGTTCTCTCCAGTTCATTTTACCGTGTGTGTAATTCGAGCAATTCT 1140  
DB 1192 CCTGCGGAGATGTTCTCTCCAGTTCATTTTACCGTGTGTGTAATTCGAGCAATTCT 1251  
QY 1141 AAAAGGCAATTTTTCGAGCCCTTGTGACTATACAGTACAGTGTCTGAAAAGCGCAG 1200  
DB 1252 AAAAGGCAATTTTTCGAGCCCTTGTGACTATACAGTACAGTGTCTGAAAAGCGCAG 1311  
QY 1201 CTACAGAGCTGTGCACTAAACAAAGGCGCAGCGATTATAGCCGCTTGTACAGATGCC 1260  
DB 1312 CTACAGAGCTGTGCACTAAACAAAGGCGCAGCGATTATAGCCGCTTGTACAGATGCC 1371  
QY 1261 TGTGCGCTGTGTGATCTCTCTGCTTCCCTTCCGACAGCCACCTCAAGCTC 1320  
DB 1372 TGTGCGCTGTGTGATCTCTCTGCTTCCCTTCCGACAGCCACCTCAAGCTC 1431  
QY 1321 CTGCTCGAATCTTCTCTAACTTCAACCCAGACCATATTTCTGTGCAAGCTCAAGTTTA 1380  
DB 1432 CTGCTCGAATCTTCTCTAACTTCAACCCAGACCATATTTCTGTGCAAGCTCAAGTTTA 1491  
QY 1381 TTTCACCCAGAAAAGCTTCATTTTGTCTTCAACATTTGATGATTTCTGTCTACGCCACA 1440  
DB 1492 TTTCACCCAGAAAAGCTTCATTTTGTCTTCAACATTTGATGATTTCTGTCTACGCCACA 1551  
QY 1441 ACAAGGTTCTGCGGAGAGGAGTATGTACAGGCTGCGCTTGTGTGCTTCAAGTT 1500  
DB 1552 ACAAGGTTCTGCGGAGAGGAGTATGTACAGGCTGCGCTTGTGTGCTTCAAGTT 1611  
QY 1501 CTTCAGCAAAATATCATGATCCCATGAAACACGCGGAAAGCCCTGCTCTTAAGATA 1560  
DB 1612 CTTCAGCAAAATATCATGATCCCATGAAACACGCGGAAAGCCCTGCTCTTAAGATA 1671  
QY 1561 TCCATCTCTCTCGAACAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
DB 1672 TCCATCTCTCTCGAACAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1731  
QY 1621 ATATATGAGGTTCCAGAACCGGACATAGCCCGTTTATTTGAGGTTCTTACAAATAGAG 1680  
DB 1732 ATATATGAGGTTCCAGAACCGGACATAGCCCGTTTATTTGAGGTTCTTACAAATAGAG 1791  
QY 1681 AAATCTCAAGAACCAACCCAGATGAAATTTTGAACAAATGTGTTTGTGCTGC 1740  
DB 1792 AAATCTCAAGAACCAACCCAGATGAAATTTTGAACAAATGTGTTTGTGCTGC 1851  
QY 1741 AGGATATAGGATAGGATATCTATTCAGAAAAGAGCTCAGACATTTCTTAAAGCATGG 1800  
DB 1852 AGGATATAGGATAGGATATCTATTCAGAAAAGAGCTCAGACATTTCTTAAAGCATGG 1911  
QY 1801 ATCTTAATCTATCTAAAGGTTTCTTCTCAAGAGATGCTCTGTGTGGAGAGAGAGCC 1860  
DB 1912 ATCTTAATCTATCTAAAGGTTTCTTCTCAAGAGATGCTCTGTGTGGAGAGAGAGCC 1971  
QY 1861 CCAGCAAAATATATGTAACAACAACATCCAGCTTACGCGCAGAGCTGCGAGATCTC 1920  
DB 1972 CCAGCAAAATATATGTAACAACAACATCCAGCTTACGCGCAGAGCTGCGAGATCTC 2031  
QY 1921 CTCAGAGAAAGCGCATATTTATGTGTGTGAGATGCAAAAGATATGCGCAAGAGATGA 1980

DB 2032 CTCAGAGAAAGCGCATATTTATGTGTGAGATGCAAAAGATATGCGCAAGATGA 2091  
QY 1981 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGTTGAGTTGAAAACTAGAAAGATG 2040  
DB 2092 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGTTGAGTTGAAAACTAGAAAGATG 2151  
QY 2041 AAAACCTGCGCACTTAAAGAGAAAGAAACGCTACCTTCAAGATATTTGTCTATA 2097  
DB 2152 AAAACCTGCGCACTTAAAGAGAAAGAAACGCTACCTTCAAGATATTTGTCTATA 2208

RESULT 7  
ADM43212  
ID ADM43212 standard; cDNA; 2094 BP.  
XX  
AC ADM43212;  
XX  
DT 03-JUN-2004 (first entry)  
XX  
DE Human methionine synthase reductase CDS G110A variant.  
XX  
KW Human; ss; Methionine synthase reductase polypeptide; HsMTR; cancer;  
KM cardiovascular disease; neural tube defect; hyperhomocysteinemia;  
XX chromosome 5p15.2-p15.3; SNP; single nucleotide polymorphism.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..2094  
FT /\*tag= a  
FT /product= "HsMTR"  
FT /partial  
FT /note= "No stop codon shown"  
FT variation replace(66,A)  
FT /\*tag= b  
FT /standard\_name= "Single\_nucleotide\_polymorphism"  
FT variation replace(110,G)  
FT /\*tag= c  
FT /standard\_name= "Single\_nucleotide\_polymorphism"  
XX  
PN US2003082676-A1.  
XX  
PD 01-MAY-2003.  
XX  
PF 10-AUG-1999; 99US-00371347.  
XX  
PR 16-JAN-1998; 98US-0071622P.  
PR 15-JAN-1999; 99US-00232028.  
XX  
PA (GRAV/) GRAVEL R A.  
PA (ROZEN/) ROZEN R.  
PA (LECLER/) LECLERC D.  
PA (WILS/) WILSON A.  
PA (ROSE/) ROSENBLATT D.  
XX  
PI Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;  
XX WPI; 2003-576610/54.  
XX  
DR P-FSDB; ADM43213.  
XX  
PT New substantially pure nucleic acid encoding a mammalian methionine  
XX synthase reductase polypeptide, useful for diagnosing, preventing or  
PT treating conditions associated with altered methionine synthase activity,  
XX e.g. cancer.  
XX  
PS Disclosure; SEQ ID NO 43; 26pp; English.  
XX  
XX The invention relates to a substantially pure nucleic acid that encodes a  
XX mammalian methionine synthase reductase polypeptide, HsMTR, or that  
XX hybridizes at high stringency to a nucleic acid appearing as ADM43208 or  
XX ADM43209. Also included are a non-human animal where one or both genetic  
XX alleles encoding the methionine synthase reductase polypeptide are



CC mutated, an antibody that specifically binds the above methionine  
CC synthase reductase polypeptide, a method of detecting the presence of the  
CC above polypeptide, a method for detecting sequence variants for  
CC methionine synthase reductase in a mammal, methods of treating or  
CC preventing cancer (or cardiovascular disease or neural tube defects) in a  
CC subject, methods of screening for a compound that modulates methionine  
CC synthase reductase biological activity and a method for detecting an  
CC increased risk of developing a neural tube defect in a mammalian embryo  
CC or fetus. The nucleic acid is useful in diagnosing, preventing or  
CC treating conditions associated with altered methionine synthase activity,  
CC such as cancer, cardiovascular disease or neural tube defects, or in  
CC screening for a compound that modulates methionine synthase reductase  
CC biological activity. Naturally occurring variants of the polypeptide are  
CC also associated with hyperhomocysteinemia. The gene for HsmTR is  
CC located on chromosome 5p15.2-p15.3. The present sequence is the coding  
CC sequence of a variant human hsmTR cDNA.

XX Sequence 2094 BP; 592 A; 489 C; 480 G; 533 T; 0 U; 0 Other;

Query Match 95.0%; Score 1992; DB 11; Length 2094;

Best Local Similarity 99.9%; Pred. No. 0; Mismatches 2; Indels 0; Gaps 0;

Matches 2092; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATGAGAGGTTTCTGTTACTATATGCTACACAGCAGGAGCAGGCAAGGCCATCGCAGAA 60
DB 1 ATGAGAGGTTTCTGTTACTATATGCTACACAGCAGGAGCAGGCAAGGCCATCGCAGAA 60
QY 61 GAAATATGTAGACACCTGTGTACATGATATTTCTGCAGATCTTCACTATATTAAGTAA 120
DB 61 GAAATGTGTAGACACCTGTGTACATGATATTTCTGCAGATCTTCACTATATTAAGTAA 120
QY 121 TCCGATTAATATGACCTTAAACCCGAAACAGCTCCTGTGTGTGTGTCTTACACAG 180
DB 121 TCCGATTAATATGACCTTAAACCCGAAACAGCTCCTGTGTGTGTGTCTTACACAG 180
QY 181 GGCAACCGAGACCCACCCGACACAGCCGCAAGTTGTTAAGAAATACAGAACCAACA 240
DB 181 GGCAACCGAGACCCACCCGACACAGCCGCAAGTTGTTAAGAAATACAGAACCAACA 240
QY 241 CTGCCGGTATTTCTTTGCTCACTGCGGATGAGTTACTGAGTCTCGTATTCAGAA 300
DB 241 CTGCCGGTATTTCTTTGCTCACTGCGGATGAGTTACTGAGTCTCGTATTCAGAA 300
QY 301 TTACACTACTTTTGGCAATGGGGGAGATTAATGAATTAACGACTTAAAGCTTGGAGCC 360
DB 301 TTACACTACTTTTGGCAATGGGGGAGATTAATGAATTAACGACTTAAAGCTTGGAGCC 360
QY 361 CGGCAATTTCTATGACACTGGAACATGAGATGACTGTGAGTTTAAACCTTGGTTGAG 420
DB 361 CGGCAATTTCTATGACACTGGAACATGAGATGACTGTGAGTTTAAACCTTGGTTGAG 420
QY 421 CCGTGATGCTGTGACCTGTGGCCAGCCCTCAGAAACATTTTAAAGTCAAGCAGAGACA 480
DB 421 CCGTGATGCTGTGACCTGTGGCCAGCCCTCAGAAACATTTTAAAGTCAAGCAGAGACA 480
QY 481 GAGGAGATTAAGTGGCGCACTCCGGTGGCATCACTGCACTCTTGAAGACAGACCTTGG 540
DB 481 GAGGAGATTAAGTGGCGCACTCCGGTGGCATCACTGCACTCTTGAAGACAGACCTTGG 540
QY 541 AAGTCAGAGCTGTACACATTTGAATCTCAAGTGAAGCTTCTGAGATTCAGATTCAGGA 600
DB 541 AAGTCAGAGCTGTACACATTTGAATCTCAAGTGAAGCTTCTGAGATTCAGATTCAGGA 600
QY 601 AGAAGAGATTTCTGAGGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATTCATGTTGA 660
DB 601 AGAAGAGATTTCTGAGGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATTCATGTTGA 660
QY 661 ATGGAAGACTTAAAGTCTCACTTAACCGTTGGTACCCCACTCTCAAGAGCTCTCTG 720
DB 661 ATGGAAGACTTAAAGTCTCACTTAACCGTTGGTACCCCACTCTCAAGAGCTCTCTG 720
QY 721 AATATTCGTGTTAATCCCAAGATATTTTACAGGTATCTGAGAGAGTCTTGGCCAG 780
DB 721 AATATTCGTGTTAATCCCAAGATATTTTACAGGTATCTGAGAGAGTCTTGGCCAG 780
```

```
DB 721 AATATTCGTGTTAATCCCAAGATATTTTACAGGTATCTGAGAGAGTCTTGGCCAG 780
QY 781 GAGGAAGCCAGATATCTGTGACTTCAGAGATCCAGTTTCAAGTGCATTTCAAG 840
DB 781 GAGGAAGCCAGATATCTGTGACTTCAGAGATCCAGTTTCAAGTGCATTTCAAG 840
QY 841 GCAATTCACCTACTACAGATGATGCAATTAACCACTGCTGCTGATAGATTTGACAT 900
DB 841 GCAATTCACCTACTACAGATGATGCAATTAACCACTGCTGCTGATAGATTTGACAT 900
QY 901 TCAATACAGACTTTTCTATCACTGAGATGCTTCAAGCTGATCTGCTTACAGAT 960
DB 901 TCAATACAGACTTTTCTATCACTGAGATGCTTCAAGCTGATCTGCTTACAGAT 960
QY 961 GATTCAGAGTACAAAGCTTCTCAAGATGCTGAGCTTGAAGATTAAGAGAGACCTGC 1020
DB 961 GATTCAGAGTACAAAGCTTCTCAAGATGCTGAGCTTGAAGATTAAGAGAGACCTGC 1020
QY 1021 GTCTTTTGAATAATTAAGCAGACACAAAGAAAGAGCTTACCTTACCCAGCATATA 1080
DB 1021 GTCTTTTGAATAATTAAGCAGACACAAAGAAAGAGCTTACCTTACCCAGCATATA 1080
QY 1081 CCGCGGAGATGTTCTCTCCAGTTCAATTTTACCTGAGTCTTGAATCCGAGCAATTCCT 1140
DB 1081 CCGCGGAGATGTTCTCTCCAGTTCAATTTTACCTGAGTCTTGAATCCGAGCAATTCCT 1140
QY 1141 AAAAAAGCAATTTTTCGAGCCCTTGTGACATATACAGGAGAGTCTGAAGAGCCAG 1200
DB 1141 AAAAAAGCAATTTTTCGAGCCCTTGTGACATATACAGGAGAGTCTGAAGAGCCAG 1200
QY 1201 CTACAGAGCTGTGAGTAAACAGAGGAGCCGATTAATAGCCGTTTGAAGAGTGC 1260
DB 1201 CTACAGAGCTGTGAGTAAACAGAGGAGCCGATTAATAGCCGTTTGAAGAGTGC 1260
QY 1261 TGTGCTGTGTGTGATCTCTCTCTGCTTCTCTTCTTCTGCAAGCAGCACTCACTCTC 1320
DB 1261 TGTGCTGTGTGTGATCTCTCTCTGCTTCTCTTCTTCTGCAAGCAGCACTCACTCTC 1320
QY 1321 CTGCTCGAATCTTCTTAACTTGAACCTTGAACCAATATGCTGAGAGCTCAAGTTA 1380
DB 1321 CTGCTCGAATCTTCTTAACTTGAACCTTGAACCAATATGCTGAGAGCTCAAGTTA 1380
QY 1381 TTTCAACCCAGAAACCTCAATTTGTCTTCAATTTGGAATTTCTGCTACCTGACCA 1440
DB 1381 TTTCAACCCAGAAACCTCAATTTGTCTTCAATTTGGAATTTCTGCTACCTGACCA 1440
QY 1441 ACAGAGTTCTGCGAAGGAGATGATGACAGCTGAGCTGCTGTGTGCTTCACTT 1500
DB 1441 ACAGAGTTCTGCGAAGGAGATGATGATGACAGCTGAGCTGCTGTGTGCTTCACTT 1500
QY 1501 CTTGACCCAAATATGATGATCCATGAAACAGCGGGAAGCCCTGCTCTTAAGATA 1560
DB 1501 CTTGACCCAAATATGATGATCCATGAAACAGCGGGAAGCCCTGCTCTTAAGATA 1560
QY 1561 TCCATCTCTCTGGAACAAATAATCTTCCATTAACAGATGACCCCTCAATCCCATC 1620
DB 1561 TCCATCTCTCTGGAACAAATAATCTTCCATTAACAGATGACCCCTCAATCCCATC 1620
QY 1621 ATTAATGAGGTTCTGAGAACCCGCAATAGCCCGTTTATTTGGTCTCTACATAGAG 1680
DB 1621 ATTAATGAGGTTCTGAGAACCCGCAATAGCCCGTTTATTTGGTCTCTACATAGAG 1680
QY 1681 AAATCTCAAGAACCAACCCAGATGGAATTTTGAAGCAATGTGTTTTTGGCTGC 1740
DB 1681 AAATCTCAAGAACCAACCCAGATGGAATTTTGAAGCAATGTGTTTTTGGCTGC 1740
QY 1741 AGGCAATTAAGGATTAATCTAATTCAGAAAGAGCTCAGACATTTCTTAAGACAG 1800
DB 1741 AGGCAATTAAGGATTAATCTAATTCAGAAAGAGCTCAGACATTTCTTAAGACAG 1800
QY 1801 ATCTTAATCTCAATTAAGGTTTCTTCTCAAGAGATGCTCTGTGTGGAGAGAGAGCC 1860
DB 1801 ATCTTAATCTCAATTAAGGTTTCTTCTCAAGAGATGCTCTGTGTGGAGAGAGAGCC 1860
```

QY 1861 CCAGCAAGTATGTACAAACAATCATTCAGCTTCATGCGCAGAGAGATTCCTC 1920  
DB 1861 CCAGCAAGTATGTACAAACAATCATTCAGCTTCATGCGCAGAGAGATTCCTC 1920  
QY 1921 CTCAGAGAGACGGCCATTTATGTTGTGTGAGATGCAAGATATGCGCAAGATGTA 1980  
DB 1921 CTCAGAGAGACGGCCATTTATGTTGTGTGAGATGCAAGATATGCGCAAGATGTA 1980  
QY 1981 CATGATGCTCTGTGCAATATATAGCAAGAGGTTGAGTTGAAAACTAGAGCAATG 2040  
DB 1981 CATGATGCTCTGTGCAATATATAGCAAGAGGTTGAGTTGAAAACTAGAGCAATG 2040  
QY 2041 AAAACCTGGCCACTTTAAAAAGAAAGAAAGCTACTTCAAGATATTGTGCA 2094  
DB 2041 AAAACCTGGCCACTTTAAAAAGAAAGAAAGCTACTTCAAGATATTGTGCA 2094

RESULT 8  
ID AAA58935 standard; DNA; 3259 BP.  
XX  
AC AAA58935;  
XX  
DT 07-NOV-2000 (first entry)  
XX  
DE DNA encoding a human methionine synthase reductase polypeptide.  
XX  
KM Human: methionine synthase reductase; MTRR; cancer;  
KM cardiovascular disease; Down's Syndrome; neural tube defect;  
KM premature coronary artery disease; 88.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 80..2176  
FT /tag= a  
FT /product= "methionine synthase reductase"  
XX  
PN WO20042196-A2.  
XX  
PD 20-JUL-2000.  
XX  
PE 14-JAN-2000; 2000WO-IB000209.  
XX  
PR 15-JAN-1999; 99US-00232028.  
XX 10-AUG-1999; 99US-00371347.  
XX  
PA (UIMC-) UNIV MCGILL.  
XX  
PI Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;  
XX  
XX WPI: 2000-466131/40.  
XX P-PSDB; AAB07591.  
XX  
PT Mammalian methionine synthase reductase nucleic acid used for detecting  
PT an increased risk of developing a neural tube defect, Down's Syndrome or  
PT cardiovascular disease in a mammalian embryo or fetus.  
XX  
PS Claim 3; Fig 3; 85pp; English.  
XX  
XX The present sequence encodes a human methionine synthase reductase (MTRR)  
XX polypeptide. Inhibitors of MTRR polypeptide and polynucleotide are used  
XX for treating or preventing cancer, cardiovascular disease, Down's  
XX Syndrome or neural tube defects in a subject. The cardiovascular disease  
XX is premature coronary artery disease. The compounds are detected by  
XX methods which screen for modulators of MTRR biological activity. MTRR  
XX polypeptide or nucleic acid is examined for the presence of a  
XX polymorphism in the parents or the embryo or foetus, and the information  
XX used for detecting an increased risk of an embryo or foetus developing  
XX cancer, cardiovascular disease, Down's Syndrome or neural tube defects  
XX  
SQ Sequence 3259 BP; 944 A; 706 C; 663 G; 946 T; 0 U; 0 Other;

Query Match 92.7%; Score 1944; DB 3; Length 3259;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2094; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGAGGAGGTTCTGTGTAATATATGCTACACAGCAGGAGCAGCAAGGCCATCGCAGAA 60  
DB 80 ATGAGGAGGTTCTGTGTAATATATGCTACACAGCAGGAGCAGCAAGGCCATCGCAGAA 139  
QY 61 GAAATATGTAGACCAAGCTGTGTGATCATGATTTTCTGACATCTTCACTGATTAAGTGA 120  
DB 140 GAAATATGTAGACCAAGCTGTGTGATCATGATTTTCTGACATCTTCACTGATTAAGTGA 199  
QY 121 TCCGATTAAGTATGACTTAAAAACCGAAACAGCTCTCTGTGTGTGTGTTTCTACCAAG 180  
DB 200 TCCGATTAAGTATGACTTAAAAACCGAAACAGCTCTCTGTGTGTGTGTTTCTACCAAG 259  
QY 181 GGCACCGGAGACCCACCGCAGACAGCCCGCAAGTTGTTAAGAAATACAGAACCAACA 240  
DB 260 GGCACCGGAGACCCACCGCAGACAGCCCGCAAGTTGTTAAGAAATACAGAACCAACA 319  
QY 241 CTGCCGGTGAATTTCTTGTCTCACTGCGGTATGGTTACTGGTCTCGGTATTCAGAA 300  
DB 320 CTGCCGGTGAATTTCTTGTCTCACTGCGGTATGGTTACTGGTCTCGGTATTCAGAA 379  
QY 301 TACACTTACTTTTGCATATGGGGGAAATATTTGATTAACGACTTCAAGAGCTTGGAGCC 360  
DB 380 TACACTTACTTTTGCATATGGGGGAAATATTTGATTAACGACTTCAAGAGCTTGGAGCC 439  
QY 361 CGGCATTTCTATGACACTGAGACATGAGATGACTGTGTGTTAGAACTTGTGTTGAG 420  
DB 440 CGGCATTTCTATGACACTGAGACATGAGATGACTGTGTGTTAGAACTTGTGTTGAG 499  
QY 421 CCGTGAATTCCTGACTCTGCGCAGCCCTCAGAAAGCAATTTAGTCAAGCAGAGACAA 480  
DB 500 CCGTGAATTCCTGACTCTGCGCAGCCCTCAGAAAGCAATTTAGTCAAGCAGAGACAA 559  
QY 481 GAGGAATTAAGTGGCGCACTCCCGGTGGCATCACTGCACTCTTGAAGACAGACCTTGTG 540  
DB 560 GAGGAATTAAGTGGCGCACTCCCGGTGGCATCACTGCACTCTTGAAGACAGACCTTGTG 619  
QY 541 AAGTCAGAGCTGTACACATTTGAATCTCAAGTGCAGCTTCTGAGATTCGATATTCAGGA 600  
DB 620 AAGTCAGAGCTGTACACATTTGAATCTCAAGTGCAGCTTCTGAGATTCGATATTCAGGA 679  
QY 601 AGAAGGATTCGAGGTTTGAAGCAAAATGCAGTGAACAGCAACCAATCCAAATGTTGA 660  
DB 680 AGAAGGATTCGAGGTTTGAAGCAAAATGCAGTGAACAGCAACCAATCCAAATGTTGA 729  
QY 661 ATTGAAGACTTTGAGTCTTCATCTTACCCGTTGGGTACCCCACTTTCACAGGCTCTCTG 720  
DB 740 ATTGAAGACTTTGAGTCTTCATCTTACCCGTTGGGTACCCCACTTTCACAGGCTCTCTG 799  
QY 721 AATATTCCTGGTTTACCCCGAGAAATTTTACAGGTATCATCTGCAGAGAGTCTCTGGCAG 780  
DB 800 AATATTCCTGGTTTACCCCGAGAAATTTTACAGGTATCATCTGCAGAGAGTCTCTGGCAG 859  
QY 781 GAGGAAAGCCAAATATCTGACTTGCAGCAGATCCAGTTTTCAGTGCCTCAATTCAGAG 840  
DB 860 GAGGAAAGCCAAATATCTGACTTGCAGCAGATCCAGTTTTCAGTGCCTCAATTCAGAG 919  
QY 841 GCAATTCACCTTACATGCAATGATGCCATTAACCACTCTGCTGTGTAATTTGACATT 900  
DB 920 GCAATTCACCTTACATGCAATGATGCCATTAACCACTCTGCTGTGTAATTTGACATT 979  
QY 901 TCAATATACAGACTTTCTTATGAGCCTGAGATGCTTCAAGGATCTGCTCAACAGT 960  
DB 980 TCAATATACAGACTTTCTTATGAGCCTGAGATGCTTCAAGGATCTGCTCAACAGT 1039  
QY 961 GATTTGAGGTACAAAGCCTTATCCAAAGACTGAGCTTGAAGATTAAGAGAGCACTGC 1020  
DB 1040 GATTTGAGGTACAAAGCCTTATCCAAAGACTGAGCTTGAAGATTAAGAGAGCACTGC 1099

1021 GTCTTTGAAAAATAGGACAGACAAAGAAAGAGTACTTATCCCGCATATA 1080  
 1100 GTCTTTGAAAAATAGGACAGACAAAGAAAGAGTACTTATCCCGCATATA 1159  
 1081 CCTGGGGAGTGTCTCTCAATTATTTTACCTGTGTCTGAATCCGAGCAATTCCT 1140  
 1160 CCGCGGGAGTGTCTCTCAATTATTTTACCTGTGTCTGAATCCGAGCAATTCCT 1219  
 1141 AAAAGGCGATTTTGGGAGCCCTTGGGACTATACAGAGACAGTGTGAAAAAGCCAG 1200  
 1220 AAAAGGCGATTTTGGGAGCCCTTGGGACTATACAGAGACAGTGTGAAAAAGCCAG 1279  
 1201 CTACAGAGCTGTGAGTAAACAGAGGGGAGCCGATTAATAGCCGCTTGTACAGATGCC 1260  
 1280 CTACAGAGCTGTGAGTAAACAGAGGGGAGCCGATTAATAGCCGCTTGTACAGATGCC 1339  
 1261 TGTGCTGTGTGTGATCTCTCTGCTTTCCTTTCGACGACACACTCACTCTC 1320  
 1340 TGTGCTGTGTGTGATCTCTCTGCTTTCCTTTCGACGACACACTCACTCTC 1399  
 1321 CTGCTCGAATCTTCTTAACTTCAACCCAGACATATTCGTGTGCAAGCTCAAGTTA 1380  
 1400 CTGCTCGAATCTTCTTAACTTCAACCCAGACATATTCGTGTGCAAGCTCAAGTTA 1459  
 1381 TTTCAACCCAGAAAGCTCAATTTTGTCTTCAACATTTGTGAATTTCTGTCTACTGCA 1440  
 1460 TTTCAACCCAGAAAGCTCAATTTTGTCTTCAACATTTGTGAATTTCTGTCTACTGCA 1519  
 1441 ACAGAGTTCTGCGAAGGAGATGTATACAGCTGCGCTGCTTGTGTGCTTCAAGTT 1500  
 1520 ACAGAGTTCTGCGAAGGAGATGTATACAGCTGCGCTGCTTGTGTGCTTCAAGTT 1579  
 1501 CTTGAGCCAAACATATGATCCCATGAAAGAGGGGAAAGCCCTGCTCTTAAGATA 1560  
 1580 CTTGAGCCAAACATATGATCCCATGAAAGAGGGGAAAGCCCTGCTCTTAAGATA 1639  
 1561 TCCATCTCTCTGGAACAACAATTTCTTCACTTACAGATAGCCCTCAATCCCATC 1620  
 1640 TCCATCTCTCTGGAACAACAATTTCTTCACTTACAGATAGCCCTCAATCCCATC 1699  
 1621 ATATATGTGGGTCCAGAAACCGGATAGCCCGTTTATTTGGGTTCTTACAACTAGAG 1680  
 1700 ATATATGTGGGTCCAGAAACCGGATAGCCCGTTTATTTGGGTTCTTACAACTAGAG 1759  
 1681 AAATCTCAAGAAACAACCCAGATGGAATTTTGGAGCAATGTGTGTTTTTGGCTGC 1740  
 1760 AAATCTCAAGAAACAACCCAGATGGAATTTTGGAGCAATGTGTGTTTTTGGCTGC 1819  
 1741 AGGATAGGATAGGATATATCAATCAGAAAAAGCTGAGACATTTCTTAAGCATGG 1800  
 1820 AGGATAGGATAGGATATATCAATCAGAAAAAGCTGAGACATTTCTTAAGCATGG 1879  
 1801 ATCTTAATCATTAAGGTTTCTTCTCAAGATGCTCTGTGTGGGAGGAGAGCC 1860  
 1880 ATCTTAATCATTAAGGTTTCTTCTCAAGATGCTCTGTGTGGGAGGAGAGCC 1939  
 1861 CCAGCAAGTATATACAGACATCCAGCTTCAATGCGCAGAGAGTGTGCAATCTC 1920  
 1940 CCAGCAAGTATATACAGACATCCAGCTTCAATGCGCAGAGAGTGTGCAATCTC 1999  
 1921 CTCACAGAGAGGCGCATATTTATGTGTGAGATGCAAAAGAAATAGCCAGATGTA 1980  
 2000 CTCACAGAGAGGCGCATATTTATGTGTGAGATGCAAAAGAAATAGCCAGATGTA 2059  
 1981 CATGATGCCCTTGTGCAATATATAGCAAGAGGTTGAGTTGAAAACTAGAGCAATG 2040  
 2060 CATGATGCCCTTGTGCAATATATAGCAAGAGGTTGAGTTGAAAACTAGAGCAATG 2119  
 2041 AAAACCTGGCCCTTTAAAGAAAGAAAAAGCTTACCTTCAAGATTTTGTGCTATA 2097  
 2120 AAAACCTGGCCCTTTAAAGAAAGAAAAAGCTTACCTTCAAGATTTTGTGCTATA 2176

RESULT 9  
 ADM43216  
 ID ADM43216 standard; cDNA; 2091 BP.  
 XX  
 AC ADM43216;  
 XX  
 DT 03-JUN-2004 (first entry)  
 XX  
 DE Human methionine synthase reductase CDS del 1726-1728 variant.  
 XX  
 KW Human; ss; Methionine synthase reductase polypeptide; HmTRR; cancer;  
 KW cardiovascular disease; neural tube defect; hyperhomocysteinemia;  
 KM chromosome 5p15.2-p15.3; SNP; single nucleotide polymorphism.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT 1..2091  
 FT /tag= a  
 FT /product= "hmTRRdelR59"  
 FT /partial  
 FT /note= "No stop codon shown"  
 FT replace(66,A)  
 FT /tag= b  
 FT /standard\_name= "single\_nucleotide\_polymorphism"  
 FT replace(110,A)  
 FT /tag= c  
 FT /standard\_name= "single\_nucleotide\_polymorphism"  
 FT replace(1726,TTGT)  
 FT /tag= d  
 PN US2003082676-A1.  
 PD 01-MAY-2003.  
 XX  
 XX 10-AUG-1999; 99US-00371347.  
 XX  
 PR 16-JAN-1998; 98US-0071622P.  
 PR 15-JAN-1999; 99US-00232028.  
 XX  
 XX (GRAV/) GRAVEL R A.  
 PA (ROZE/) ROZEN R.  
 PA (LECL/) LECLERC D.  
 PA (WILS/) WILSON A.  
 PA (ROSE/) ROSENBLATT D.  
 PI Grave! RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;  
 DR WPI: 2003-576610/54.  
 DR P-PsDB; ADM43217.  
 XX  
 PT New substantially pure nucleic acid encoding a mammalian methionine  
 PT synthase reductase polypeptide, useful for diagnosing, preventing or  
 PT treating conditions associated with altered methionine synthase activity,  
 PT e.g. cancer.  
 PT  
 XX  
 PS Disclosure; SEQ ID NO 45; 26bp; English.  
 XX  
 XX The invention relates to a substantially pure nucleic acid that encodes a  
 CC mammalian methionine synthase reductase polypeptide, HmTRR, or that  
 CC hybridizes at high stringency to a nucleic acid appearing as ADM43208 or  
 CC ADM43209. Also included are a non-human animal where one or both genetic  
 CC alleles encoding the methionine synthase reductase polypeptide are  
 CC mutated, an antibody that specifically binds the above methionine  
 CC synthase reductase polypeptide, a method of detecting the presence of the  
 CC above polypeptide, a method for detecting sequence variants for  
 CC methionine synthase reductase in a mammal, methods of treating or  
 CC preventing cancer (or cardiovascular disease or neural tube defects) in a  
 CC subject, methods of screening for a compound that modulates methionine  
 CC synthase reductase biological activity and a method for detecting an  
 CC increased risk of developing a neural tube defect in a mammalian embryo  
 CC or foetus. The nucleic acid is useful in diagnosing, preventing or  
 CC treating conditions associated with altered methionine synthase activity,

CC such as cancer, cardiovascular disease or neural tube defects, or in  
CC screening for a compound that modulates methionine synthase reductase  
CC biological activity. Naturally occurring variants of the polypeptide are  
CC also associated with hyperhomocysteinemia. The gene for Hmtrr is  
CC located on chromosome 5p15.2-p15.3. The present sequence is the coding  
CC sequence of a variant human hmtrr cDNA.

XX Sequence 2091 BP; 591 A; 489 C; 480 G; 531 T; 0 U; 0 Other;

Query Match 85.8%; Score 1800; DB 11; Length 2091;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2091; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

```
QY 1 ATGAGAGGTTTCTGTTACTATATGCTACACAGAGGAGACAGCAAGGCCATCGCAGAA 60
DB 1 ATGAGAGGTTTCTGTTACTATATGCTACACAGAGGAGACAGCAAGGCCATCGCAGAA 60
QY 61 GAAATATGTAGAGCAAGCTGTGTACATGAGATTTCTGCAATCTTCACTGTATTAATGAA 120
DB 61 GAAATATGTAGAGCAAGCTGTGTACATGAGATTTCTGCAATCTTCACTGTATTAATGAA 120
QY 121 TCCGATATGATGACTATAAACCAGAAACAGCTCTCTGTTGTGTGTCTTCAACAG 180
DB 121 TCCGATATGATGACTATAAACCAGAAACAGCTCTCTGTTGTGTGTCTTCAACAG 180
QY 121 TCCGATATGATGACTATAAACCAGAAACAGCTCTCTGTTGTGTGTCTTCAACAG 180
DB 121 TCCGATATGATGACTATAAACCAGAAACAGCTCTCTGTTGTGTGTCTTCAACAG 180
QY 181 GGCACCGGAGACCCACCCGACACAGCCCGCAAGTTTGAAGAAATACAGAAACAAACA 240
DB 181 GGCACCGGAGACCCACCCGACACAGCCCGCAAGTTTGAAGAAATACAGAAACAAACA 240
QY 241 CTGCGCGGTATTTCTTGTCTGACCTGTGGGATGAGGTTTCTGCGTCTCGGTATTCAGAA 300
DB 241 CTGCGCGGTATTTCTTGTCTGACCTGTGGGATGAGGTTTCTGCGTCTCGGTATTCAGAA 300
QY 301 TACACCTACTTTTGGCAATGAGGAGGAAATATGATTAAGCACTTCAAGAGCTTGAGCC 360
DB 301 TACACCTACTTTTGGCAATGAGGAGGAAATATGATTAAGCACTTCAAGAGCTTGAGCC 360
QY 361 CGGCAATTTCTATGACACTGACATGACATGACATGCTGTAGGTTTGAACCTTGTGTGAG 420
DB 361 CGGCAATTTCTATGACACTGACATGACATGACATGCTGTAGGTTTGAACCTTGTGTGAG 420
QY 421 CCGGATTTGCTGTGACCTGTGGCCAGCCCTCAGAAAGATTTTGAAGTAAACAGAGGACAA 480
DB 421 CCGGATTTGCTGTGACCTGTGGCCAGCCCTCAGAAAGATTTTGAAGTAAACAGAGGACAA 480
QY 481 GAGGAGATTAAGTGGCGCACTCCCGGTGGCAATCACTGCATCTTGAAGACAGACCTTGTG 540
DB 481 GAGGAGATTAAGTGGCGCACTCCCGGTGGCAATCACTGCATCTTGAAGACAGACCTTGTG 540
QY 541 AAGTCAAGCTGTGACATGATTAATCTCAAGTGTGAGCTTGTGAGATTCGATTCAGGA 600
DB 541 AAGTCAAGCTGTGACATGATTAATCTCAAGTGTGAGCTTGTGAGATTCGATTCAGGA 600
QY 601 AGAAGGATTTCTGAGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATCCAAATGTTGA 660
DB 601 AGAAGGATTTCTGAGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATCCAAATGTTGA 660
QY 661 ATTGAAGATTTGAGTCTCACTTACCGTTCGTTACCCCACTCTCAAGCCTCTCTG 720
DB 661 ATTGAAGATTTGAGTCTCACTTACCGTTCGTTACCCCACTCTCAAGCCTCTCTG 720
QY 721 AATATTCCTGTGTTTACCCCGAATATTTTCAAGTACATCTGACAGAGTCTCTGGCCAG 780
DB 721 AATATTCCTGTGTTTACCCCGAATATTTTCAAGTACATCTGACAGAGTCTCTGGCCAG 780
QY 781 GAGGAAAGCCAGATATCTGTGACTTTCAGCAAGTCCAGTTTTCAGGCAATTTCAAG 840
DB 781 GAGGAAAGCCAGATATCTGTGACTTTCAGCAAGTCCAGTTTTCAGGCAATTTCAAG 840
QY 841 GCAGTTCAACTTACTAGCAATGATGCAATAAACAACCTGCTGTGATGAAATTTGCAATT 900
DB 841 GCAGTTCAACTTACTAGCAATGATGCAATAAACAACCTGCTGTGATGAAATTTGCAATT 900
```

```
QY 901 TCAATATCAGACTTTTCTATACAGCTTGAGATGCTTACAGCGTATCTGCCCTAACATG 960
DB 901 TCAATATCAGACTTTTCTATACAGCTTGAGATGCTTACAGCGTATCTGCCCTAACATG 960
QY 961 GATTTGAGGTTCAAAGCCATCTCCAAAGCTGACCTTGAGATTAAGAGAGACCTGC 1020
DB 961 GATTTGAGGTTCAAAGCCATCTCCAAAGCTGACCTTGAGATTAAGAGAGACCTGC 1020
QY 1021 GTCTTTTGAATAAATGAAGCAGACACAAAGAAAGAGCTACCTTACCCAGCATATA 1080
DB 1021 GTCTTTTGAATAAATGAAGCAGACACAAAGAAAGAGCTACCTTACCCAGCATATA 1080
QY 1081 CCGCGGAGATGTTCTCTCCAGTTCAATTTTCTGCTGCTTGAATCCGAGAAATTCCT 1140
DB 1081 CCGCGGAGATGTTCTCTCCAGTTCAATTTTCTGCTGCTTGAATCCGAGAAATTCCT 1140
QY 1141 AAAAAGCAATTTTGGAGCCCTGTGGACTATACAGTGAACAGTCTGTAAGAGCCAGG 1200
DB 1141 AAAAAGCAATTTTGGAGCCCTGTGGACTATACAGTGAACAGTCTGTAAGAGCCAGG 1200
QY 1201 CTACAGAGCTGTGACATTAACAAAGGAGGACCCGATTAATAGCCCTTGTACAGATGCC 1260
DB 1201 CTACAGAGCTGTGACATTAACAAAGGAGGACCCGATTAATAGCCCTTGTACAGATGCC 1260
QY 1261 TGTGCTGCTTGTGATCTCTCTGCTGCTTTCCTTTCGACGACACATCACTCTC 1320
DB 1261 TGTGCTGCTTGTGATCTCTCTGCTGCTTTCCTTTCGACGACACATCACTCTC 1320
QY 1321 CTGCTGAAATCTTCTTAACTTCAACCCAGACCAATAATTCGTGACAGCTCAAGTTTA 1380
DB 1321 CTGCTGAAATCTTCTTAACTTCAACCCAGACCAATAATTCGTGACAGCTCAAGTTTA 1380
QY 1381 TTTCAACCAGAAAGCTCCATTTTGTCTTCAACATTTGTGGAATTTGTCTACTGACACA 1440
DB 1381 TTTCAACCAGAAAGCTCCATTTTGTCTTCAACATTTGTGGAATTTGTCTACTGACACA 1440
QY 1441 ACAAGGTTCTGCGGAAAGGAGATGATACAGCTGCTGCTGCTTGTGTTGTTCAATT 1500
DB 1441 ACAAGGTTCTGCGGAAAGGAGATGATGATGATGATGATGATGATGATGATGATGAT 1500
QY 1501 CTTCAAGCAAACTATCATGATCCCATGGAAGACAGGAGGAAAGCCCTGAGCTCTAAGATA 1560
DB 1501 CTTCAAGCAAACTATCATGATCCCATGGAAGACAGGAGGAAAGCCCTGAGCTCTAAGATA 1560
QY 1561 TCCATCTCTCTGGAACAAACAAATTTTTCATTTCAAGATGACCCCTCAATCCCATC 1620
DB 1561 TCCATCTCTCTGGAACAAACAAATTTTTCATTTCAAGATGACCCCTCAATCCCATC 1620
QY 1621 ATTAATGTTGTTCAAGAAACCGGATAGCCCTTATTTGAGTTTCTTCAACATAGAGAG 1680
DB 1621 ATTAATGTTGTTCAAGAAACCGGATAGCCCTTATTTGAGTTTCTTCAACATAGAGAG 1680
QY 1681 AAATCTCAAGAAACAAACCAAGATGGAATTTTGAAGCAATGTGTGTTTGTGCTGC 1740
DB 1681 AAATCTCAAGAAACAAACCAAGATGGAATTTTGAAGCAATGTGTGTTTGTGCTGC 1740
QY 1741 AGCATTAAGATGAGATTAATCTATTCAGAAAGACTCAGACATTTCTTAAGCATGG 1800
DB 1741 AGCATTAAGATGAGATTAATCTATTCAGAAAGACTCAGACATTTCTTAAGCATGG 1800
QY 1738 AGGCAATAGGATGAGATTAATCTATTCAGAAAGACTCAGACATTTCTTAAGCATGG 1797
DB 1738 AGGCAATAGGATGAGATTAATCTATTCAGAAAGACTCAGACATTTCTTAAGCATGG 1797
QY 1801 ATCTTAATCTATTAAGGTTTCTTCTCAAGAGATGCTCTGTGGGAGAGAGAGCC 1860
DB 1801 ATCTTAATCTATTAAGGTTTCTTCTCAAGAGATGCTCTGTGGGAGAGAGAGCC 1860
QY 1798 ATCTTAATCTATTAAGGTTTCTTCTCAAGAGATGCTCTGTGGGAGAGAGAGCC 1857
DB 1798 ATCTTAATCTATTAAGGTTTCTTCTCAAGAGATGCTCTGTGGGAGAGAGAGCC 1857
QY 1861 CAGCAAGATGATTAAGACACATCCAGCTTCAAGGACAGAGGAGGAGAGATCTC 1920
DB 1861 CAGCAAGATGATTAAGACACATCCAGCTTCAAGGACAGAGGAGGAGAGATCTC 1920
QY 1858 CAGCAAGATGATTAAGACACATCCAGCTTCAAGGACAGAGGAGGAGAGATCTC 1917
DB 1858 CAGCAAGATGATTAAGACACATCCAGCTTCAAGGACAGAGGAGGAGAGATCTC 1917
QY 1921 CTCAGAGAAACGGCCATATTTATGTGTGAGATGCAAGAAATATGCAAGATGTA 1980
DB 1921 CTCAGAGAAACGGCCATATTTATGTGTGAGATGCAAGAAATATGCAAGATGTA 1980
QY 1918 CTCAGAGAAACGGCCATATTTATGTGTGAGATGCAAGAAATATGCAAGATGTA 1977
DB 1918 CTCAGAGAAACGGCCATATTTATGTGTGAGATGCAAGAAATATGCAAGATGTA 1977
QY 1981 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGGTTGAGTTGAAAACTAAGAACATG 2040
```

```
DB 1978 CATGATGCGCTTGTGCAATATATAGCAAGAGGTTGAGTTGAAAACTAGAGCAATG 2037
QY 2041 AAAACCTGGCCACTTTAAAAGAGAAAACGCTACCTTCAGAGATTTTGTCA 2094
DB 2038 AAAACCTGGCCACTTTAAAAGAGAAAACGCTACCTTCAGAGATTTTGTCA 2091

RESULT 10
ADM43214
ID ADM43214 standard; cDNA; 2091 BP.
XX
AC ADM43214;
XX
DT 03-JUN-2004 (first entry)
XX
DE Human methionine synthase reductase CDS del 1675-1678 variant.
XX
KM Human; ss; Methionine synthase reductase polypeptide; hsmTRR; cancer;
XX cardiovacular disease; neural tube defect; hyperhomocysteinemia;
XX chromosome 5p15.2-p15.3; SNP; single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..2091
FT /tag= a
FT /product= "hsmTRRdelR559"
FT /partial
FT /note= "No stop codon shown"
FT variation replace(66,A)
FT /tag= b
FT /standard_name= "Single_nucleotide_polymorphism"
FT /tag= c
FT /standard_name= "Single_nucleotide_polymorphism"
FT variation replace(1675,AGAG)
FT /tag= d
XX
PN US2003082676-A1.
XX
PD 01-MAY-2003.
XX
PP 10-AUG-1999; 99US-00371347.
XX
PR 16-JAN-1998; 98US-0071622P.
XX 15-JAN-1999; 99US-00232028.
XX
PA (GRAY/) GRAVEL R. A.
PA (ROZE/) ROZEN R.
PA (LECL/) LECLERC D.
PA (WILS/) WILSON A.
PA (ROSE/) ROSENBLATT D.
XX
PI Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;
XX
DR WPI; 2003-576610/54.
DR P-PSDB; ADM43215.
XX
PT New substantially pure nucleic acid encoding a mammalian methionine
PT synthase reductase polypeptide, useful for diagnosing, preventing or
PT treating conditions associated with altered methionine synthase activity,
PT e.g. cancer.
XX
PS Disclosure; SEQ ID NO 47; 26pp; English.
XX
CC The invention relates to a substantially pure nucleic acid that encodes a
CC mammalian methionine synthase reductase polypeptide, hsmTRR, or that
CC hybridizes at high stringency to a nucleic acid appearing as ADM43208 or
CC ADM43209. Also included are a non-human animal where one or both genetic
CC alleles encoding the methionine synthase reductase polypeptide are
CC mutated, an antibody that specifically binds the above methionine
CC synthase reductase polypeptide, a method of detecting the presence of the
```

```
CC above polypeptide, a method for detecting sequence variants for
CC methionine synthase reductase in a mammal, methods of treating or
CC preventing cancer (or cardiovascular disease or neural tube defects) in a
CC subject, methods of screening for a compound that modulates methionine
CC synthase reductase biological activity and a method for detecting an
CC increased risk of developing a neural tube defect in a mammalian embryo
CC or foetus. The nucleic acid is useful in diagnosing, preventing or
CC treating conditions associated with altered methionine synthase activity,
CC such as cancer, cardiovascular disease or neural tube defects, or in
CC screening for a compound that modulates methionine synthase reductase
CC biological activity. Naturally occurring variants of the polypeptide are
CC also associated with hyperhomocysteinemia. The gene for hsmTRR is
CC located on chromosome 5p15.2-p15.3. The present sequence is the coding
CC sequence of a variant human hsmTRR cDNA.
XX
SQ Sequence 2091 BP; 589 A; 489 C; 480 G; 533 T; 0 U; 0 Other;
XX
Query Match 85.8%; Score 1800; DB 11; Length 2091;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2090; Conservative 0; Mismatches 1; Indels 3; Gaps 1;
QY 1 ATGAGAGGTTTCTGTACTATATATGCTACAGCGAGGACAGGCAAGGCCATCGCAGAA 60
DB 1 ATGAGAGGTTTCTGTACTATATATGCTACAGCGAGGACAGGCAAGGCCATCGCAGAA 60
QY 61 GAAATATGAGACAGCTGTGTACATGATTTTTCGAGATCTTCACTGTATAGTGAA 120
DB 61 GAAATATGAGACAGCTGTGTACATGATTTTTCGAGATCTTCACTGTATAGTGAA 120
QY 121 TCCGATTAAGTATGACCTAAACCCGAAACAGCTCCTCTGTGTGTGTGTTTCTACACG 180
DB 121 TCCGATTAAGTATGACCTAAACCCGAAACAGCTCCTCTGTGTGTGTGTTTCTACACG 180
QY 181 GGCACCGAGAGACCCAGCCGACAGCCGCAAGTTGTTAAGAAATACAGAACCAACA 240
DB 181 GGCACCGAGAGACCCAGCCGACAGCCGCAAGTTGTTAAGAAATACAGAACCAACA 240
QY 241 CCGCGGTGATTTCTTGTGCTCACTGCGGTATGAGTTATCGGGTCTCGGTATTCAGAA 300
DB 241 CCGCGGTGATTTCTTGTGCTCACTGCGGTATGAGTTATCGGGTCTCGGTATTCAGAA 300
QY 301 TACACCTACTTTTGGCAATGGGGGGAATATATGATTAACGACTTCAGAGCTTGAGCC 360
DB 301 TACACCTACTTTTGGCAATGGGGGGAATATATGATTAACGACTTCAGAGCTTGAGCC 360
QY 361 CGGCAATTTCTATGACACTGACATGCAATGCACTGTGATGTTAGAACTTGTGTTGAG 420
DB 361 CGGCAATTTCTATGACACTGACATGCAATGCACTGTGATGTTAGAACTTGTGTTGAG 420
QY 421 CCGTGAATGCTGCACTGCGGCAAGCCCTCAAGAAAGCAATTTTAGGTCAGAGAGCAAA 480
DB 421 CCGTGAATGCTGCACTGCGGCAAGCCCTCAAGAAAGCAATTTTAGGTCAGAGAGCAAA 480
QY 481 GAGAGATTAAGTGGGCACTCCCGGTGATCACTGATCTTGAAGAGACAGACTTGTG 540
DB 481 GAGAGATTAAGTGGGCACTCCCGGTGATCACTGATCTTGAAGAGACAGACTTGTG 540
QY 541 AAGTCAGAGCTGCTACATTAATCTCAAGTCAAGCTTCTGAGATTCATGATTCAGGA 600
DB 541 AAGTCAGAGCTGCTACATTAATCTCAAGTCAAGCTTCTGAGATTCATGATTCAGGA 600
QY 601 AGAAGAGATTCAGAGTTTGAAGCAAAATGCAAGGAAACAGCAACCAATCCAAATGTTGA 660
DB 601 AGAAGAGATTCAGAGTTTGAAGCAAAATGCAAGGAAACAGCAACCAATCCAAATGTTGA 660
QY 661 ATTGAAGACTTGAAGTCTCACTTACCCGTTGAGTACCCCACTCTCAAGAGCTCTGTG 720
DB 661 ATTGAAGACTTGAAGTCTCACTTACCCGTTGAGTACCCCACTCTCAAGAGCTCTGTG 720
QY 721 AATATTCCTGTTTACCCCAAGATATTTTACAGGTACATCTGCAAGAGTCTTGGCCAG 780
DB 721 AATATTCCTGTTTACCCCAAGATATTTTACAGGTACATCTGCAAGAGTCTTGGCCAG 780
```





Qy	1	ATGAGGAGGTTTCGTGTAAGTAATAGTCACACAGAGGGACAGAGCAAGGCATCCGAGA	60
Db	80	ATGAGGAGGTTTCGTGTAAGTAATAGTCACACAGAGGGACAGAGCAAGGCATCCGAGA	139
Qy	61	GAATATATGAGCAAGCTGTGTATCATGSAATTTTCTGCAGATCTTCACTGTATATAGGAA	120
Db	140	GAATATGTGTAGCAAGCTGTGTATCATGSAATTTTCTGCAGATCTTCACTGTATATAGGAA	199
Qy	121	TCCGATTAAGTAATGACTTAATAAACCGAAACAGCTCTCTTGTGTGTGTTTCTACACG	180
Db	200	TCCGATTAAGTAATGACTTAATAAACCGAAACAGCTCTCTTGTGTGTGTTTCTACACG	259
Qy	181	GGCACCCGGAACCCACCCGACACAGCCCCGCAAGTTTGTAAAGAAATACGAAACCAACA	240
Db	260	GGCACCCGGAACCCACCCGACACAGCCCCGCAAGTTTGTAAAGAAATACGAAACCAACA	319
Qy	241	CTGCGCGGTGATTTCTTGTGCTCACCTGCGGATATGGGTATCTGGGTCTCGGATTTCAAA	300
Db	320	CTGCGCGGTGATTTCTTGTGCTCACCTGCGGATATGGGTATCTGGGTCTCGGATTTCAAA	379
Qy	301	TACACCTACTTCTTTCGCAATGGGGGGAGATTAATGATTAACGACTTCAAGACTTGAAGCC	360
Db	380	TACACCTACTTTCGCAATGGGGGGAGATTAATGATTAACGACTTCAAGACTTGAAGCC	439
Qy	361	CGGCATTTCTATGACACTGGAACATGCAGATGACTGTGTAGTTTGAACCTTGTGTGAG	420
Db	440	CGGCATTTCTATGACACTGGAACATGCAGATGACTGTGTAGTTTGAACCTTGTGTGAG	499
Qy	421	CCGATGATTTGTGGAAGCTGTGGCCAGCCCTCAGAAAGCAATTTAGTCAAGCAGAGACA	480
Db	500	CCGATGATTTGTGGAAGCTGTGGCCAGCCCTCAGAAAGCAATTTAGTCAAGCAGAGACA	559
Qy	481	GAGAGATTAAGTGGCGCACTCCCGGTGGCATCACTGCAATCTTGAAGACAGACTTGTG	540
Db	560	GAGAGATTAAGTGGCGCACTCCCGGTGGCATCACTGCAATCTTGAAGACAGACTTGTG	619
Qy	541	AAGTCAGAGCTGTCAACATTTGAATTCACAGTCAGCTTCTGAGATTCGATGATTCAGGA	600
Db	620	AAGTCAGAGCTGTCAACATTTGAATTCACAGTCAGCTTCTGAGATTCGATGATTCAGGA	679
Qy	601	AGAAAGATTTCTGAGTTTGAAGCAAAATGCATGAAAGCAACCAATCCAAATGTTGA	660
Db	680	AGAAAGATTTCTGAGTTTGAAGCAAAATGCATGAAAGCAACCAATCCAAATGTTGA	739
Qy	661	ATTGAAGACTTGAAGCTCTCACTTAACCGTTCGGTACCCCACTCTGACAAAGCTCTCTG	720
Db	740	ATTGAAGACTTGAAGCTCTCACTTAACCGTTCGGTACCCCACTCTGACAAAGCTCTCTG	799
Qy	721	AATATTCCTGTGTTAACCCCGAATAATTTACAGGTACATCTGCAGAGAGTCTTGTGGCAG	780
Db	800	AATATTCCTGTGTTAACCCCGAATAATTTACAGGTACATCTGCAGAGAGTCTTGTGGCAG	859
Qy	781	GAGGAAAGCCAAAGTATCTGTGACTTCAGCAGATCCAGTTTTCAGATGCCAATTTCAAG	840
Db	860	GAGGAAAGCCAAAGTATCTGTGACTTCAGCAGATCCAGTTTTCAGATGCCAATTTCAAG	919
Qy	841	GCAGTTCAACTTACTAGATATGATGCCATTAATAAACCACTGCTGTGTGAATTTGAACATT	900
Db	920	GCAGTTCAACTTACTAGATATGATGCCATTAATAAACCACTGCTGTGTGAATTTGAACATT	979
Qy	901	TCAATATCAGACTTTTCTATCAGCTGTGAGATGCTTCAAGCTGATCTGCTTACAGT	960
Db	980	TCAATATCAGACTTTTCTATCAGCTGTGAGATGCTTCAAGCTGATCTGCTTACAGT	1039
Qy	961	GATTTCTGAGGTACAAAGCTTACTCCAAAGATGCAAGCTTGAAGATTAAGAGACACTGC	1020
Db	1040	GATTTCTGAGGTACAAAGCTTACTCCAAAGATGCAAGCTTGAAGATTAAGAGACACTGC	1099
Qy	1021	GTCTTTTGAATAATTAAGCAGACACAAGAAAGAAAGAGCTTACCTTACCCAGCATATA	1080
Db	1100	GTCTTTTGAATAATTAAGCAGACACAAGAAAGAAAGAGCTTACCTTACCCAGCATATA	1159
Qy	1081	CTGTGGGAGTATCTCTCAGGTTCAATTTTATCTGTGTCTTGAATCCGAGCAATTCCT	1140

Db	1160	CTGCGGGAGTGTCTCTCCAGTTCACTTTTAACGTGGTGTGGAATCCGAGCAATTCCT	1219
Qy	1141	AAAAAGCATTTTTCGAGGCCCTTGTGACATATCCAGTGACAGTGTCTGAAMAGCGCAGG	1200
Db	1220	AAAAAGCATTTTTCGAGGCCCTTGTGACATATCCAGTGACAGTGTCTGAAMAGCGCAGG	1279
Qy	1201	CTACAGAGCGTGTGACATTAACAAAGGGGACGCCGATATATAGCCGCTTGTGTACAGATATCC	1266
Db	1280	CTACAGAGCGTGTGACATTAACAAAGGGGACGCCGATATATAGCCGCTTGTGTACAGATATCC	1339
Qy	1261	TGTGCTGTGTGTGATCTCTCTGCTTCTCTTCTTGTGCAAGCACTCAGTCTC	1320
Db	1340	TGTGCTGTGTGTGATCTCTCTGCTTCTCTCTTGTGCAAGCACTCAGTCTC	1399
Qy	1321	CTGCTCCGAACATCTTCTTAAACCTTCAACCCAGACCAATTCGTGTGCAAGCTCAAGTTTA	1380
Db	1400	CTGCTCCGAACATCTTCTTAAACCTTCAACCCAGACCAATTCGTGTGCAAGCTCAAGTTTA	1459
Qy	1381	TTTCACCCAGGAAAGCTCCATTTTGTCTTCACATTTGTGCAATTTCTGTACTACTGCACA	1440
Db	1460	TTTCACCCAGGAAAGCTCCATTTTGTCTTCACATTTGTGCAATTTCTGTACTACTGCACA	1519
Qy	1441	ACAGAGGTTCTGCGAAGGAGATATGTAACAGGCTGGCTGCTGTGTGTTCTTCAAGTT	1500
Db	1520	ACAGAGGTTCTGCGAAGGAGATATGTAACAGGCTGGCTGCTGTGTGTTCTTCAAGTT	1579
Qy	1501	CTTGACGCCAACAATACATACATGCATCTCCATGAGACAGCGGGAAAGCCCTGGCTCTTAAGATA	1560
Db	1580	CTTGACGCCAACAATACATACATGCATCTCCATGAGACAGCGGGAAAGCCCTGGCTCTTAAGATA	1639
Qy	1561	TCGATCTCTCTCGAACACACAAATTCYTCACCTTACCGATGACCCCTCAATCCCATC	1620
Db	1640	TCGATCTCTCTCGAACACACAAATTCYTCACCTTACCGATGACCCCTCAATCCCATC	1699
Qy	1621	ATATATGTGGGTCCAGGAAACCGGATAGCCCGTTTATGGGTTCTACACATATAGAGAG	1608
Db	1700	ATATATGTGGGTCCAGGAAACCGGATAGCCCGTTTATGGGTTCTACACATATAGAGAG	1756
Qy	1681	AAACTCCAAAGAACACCCAGATGAGAAATTTTGAGCAATGTGGTGTTTTTTGCGCTGC	1740
Db	1757	AAACTCCAAAGAACACCCAGATGAGAAATTTTGAGCAATGTGGTGTTTTTTGCGCTGC	1816
Qy	1741	AGGATTAAGATAGGATTAATCTAATTCAGAAAGAGCTCAGACATTTCTTAAGCATGGG	1800
Db	1817	AGGATTAAGATAGGATTAATCTAATTCAGAAAGAGCTCAGACATTTCTTAAGCATGGG	1876
Qy	1801	ATCTTAATCTCATTAAGGTTTCTCTCCAGAGATGTCTGTGTGGGAGAGAGAGACC	1866
Db	1877	ATCTTAATCTCATTAAGGTTTCTCTCCAGAGATGTCTGTGTGGGAGAGAGAGACC	1933
Qy	1861	CCAGCAAAAGTATGTACAGACACATCCAGCTTCATGGCCAGCAGGTGGCAGAAATCTCTC	1920
Db	1937	CCAGCAAAAGTATGTACAGACACATCCAGCTTCATGGCCAGCAGGTGGCAGAAATCTCTC	1996
Qy	1921	CTCCAGAGAAACGGCCATATTTATGTGTGTGGAATATGCAAGAAATATGGCCCAAGATCTA	1988
Db	1997	CTCCAGAGAAACGGCCATATTTATGTGTGTGGAATATGCAAGAAATATGGCCCAAGATCTA	2056
Qy	1981	CATGATGCTCTGTGTGCAAAATATATAGCCAAAGAGTTGGAGTTGAAAACTAGAAAGCATATG	2048
Db	2057	CATGATGCTCTGTGTGCAAAATATATAGCCAAAGAGTTGGAGTTGAAAACTAGAAAGCATATG	2116
Qy	2041	AAAAACCTGTGCACTTTAAAGAAAGAAAACGCTACCTTCAGAGATATTTGGTCAATA	2097
Db	2117	AAAAACCTGTGCACTTTAAAGAAAGAAAACGCTACCTTCAGAGATATTTGGTCAATA	2173

```

RESULT 12
AAA58976
ID   AAA58976  standard; DNA; 3255 BP
XX
AC   AAA58976;

```







Dh 520 ATTGCTGGACTTGGCCAGCCCTCAGAAAGCATTTTGGTCAAGCAGAGCAAGAGAG 579  
Qy 487 ATAAAGTGGCGACTCCCGGTGGCATCACTGCATCTTGAAGACAGACCTTGAAGTCA 546  
Dh 580 ATTAAGTGGCGACTCCCGGTGGCATCACTGCATCTTGAAGACAGACCTTGAAGTCA 639  
Qy 547 GAGCTGTACACATTTGAATCTCAAGTGGAGCTTGTGAGATTCCATGATTTCAAGAGAAAG 606  
Dh 640 GAGCTGTACACATTTGAATCTCAAGTGGAGCTTGTGAGATTCCATGATTTCAAGAGAAAG 639  
Qy 607 GATTCTGAGGTTTGAAGCAAAAATGCAGTGACAGCAACCAATCCATGTTGTAATTGAA 666  
Dh 700 GATTCTGAGGTTTGAAGCAAAAATGCAGTGACAGCAACCAATCCATGTTGTAATTGAA 759  
Qy 667 GACTTGTAGTCTCACTTACCCGTTGGTACCCCACTCTCAACAGCTCTGTGAATATT 726  
Dh 760 GACTTGTAGTCTCACTTACCCGTTGGTACCCCACTCTCAACAGCTCTGTGAATATT 819  
Qy 727 CTTGGTTTACCCCGAATATTTTACAGTACATCTGAGAGAGTCTTGGCCAGAGAGAA 786  
Dh 820 CTTGGTTTACCCCGAATATTTTACAGTACATCTGAGAGAGTCTTGGCCAGAGAGAA 879  
Qy 787 AGCCAAATATCTGTGACCTTCAAGCATCCAGTTTTCAGTGCATTTCAAAAGCAGTT 846  
Dh 880 AGCCAAATATCTGTGACCTTCAAGCATCCAGTTTTCAGTGCATTTCAAAAGCAGTT 939  
Qy 847 CAATTACTACGAATGATGCTATTAATAACAACCTCTGTGTGATGAAATTGACATTTCAAT 906  
Dh 940 CAATTACTACGAATGATGCTATTAATAACAACCTCTGTGTGATGAAATTGACATTTCAAT 999  
Qy 907 ACAGACTTTTCTCATCAGCCTGAGAGATGCTTCAAGCTGATCTGCTTACAGTGAATCT 966  
Dh 1000 ACAGACTTTTCTCATCAGCCTGAGAGATGCTTCAAGCTGATCTGCTTACAGTGAATCT 1059  
Qy 967 GAGGTACAAAGCCTACTCCAAAGACTGACGCTTGAAGATTAAGAGAGACCTGCTCTT 1026  
Dh 1060 GAGGTACAAAGCCTACTCCAAAGACTGACGCTTGAAGATTAAGAGAGACCTGCTCTT 1119  
Qy 1027 TTGAAATTAAGGACGACACAAAGAAAGAGACTTACCCTCCAGCATATACCTGCG 1086  
Dh 1120 TTGAAATTAAGGACGACACARAGAAAGAGACTTACCCTCCAGCATATACCTGCG 1179  
Qy 1087 GGAATGTTCTCTCCAGTCAATTTTAACTGCTGCTTGAATTCGAGGAATTCCTAAAG 1146  
Dh 1180 GGAATGTTCTCTCCAGTCAATTTTAACTGCTGCTTGAATTCGAGGAATTCCTAAAG 1239  
Qy 1147 GCATTTTGGAGGCTTGTGACATATACAGTGAAGTGTGAAAAGCCAGGCTACAG 1206  
Dh 1240 GCATTTTTRCGAGGCTTGTGACATATACAGTGAAGTGTGAAAAGCCAGGCTACAG 1299  
Qy 1207 GAGCTGTGACGTAAACAAGGGGAGCCGATTAATAGCCGTTTGTACGAGATGCTGTGCC 1266  
Dh 1300 GAGCTGTGACGTAAACAAGGGGAGCCGATTAATAGCCGTTTGTACGAGATGCTGTGCC 1359  
Qy 1267 TGCCTTGTGATCT 1326  
Dh 1360 TGCCTTGTGATCT 1419  
Qy 1327 GAAACATCTCTTAACTTCAACCCAGACCATATTCGTGTGACAGCTCAAGTTTATTCAC 1386  
Dh 1420 GAAACATCTCTTAACTTCAACCCAGACCATATTCGTGTGACAGCTCAAGTTTATTCAC 1479  
Qy 1387 CCAGGAAGAGTCCATTTTGTCTTCAACATTTGTGGAATTTCTGTCTACTGCGCAACAGAG 1446  
Dh 1480 CCAGGAAGAGTCCATTTTGTCTTCAACATTTGTGGAATTTCTGTCTACTGCGCAACAGAG 1539  
Qy 1447 GTTCTGGGGAAGGAGATGACAGGCTGGCTGGCTTGTGTGCTTCAAGTTCTTCAAG 1506  
Dh 1540 GTTCTGGGGAAGGAGATGACAGGCTGGCTGGCTTGTGTGCTTCAAGTTCTTCAAG 1599  
Qy 1507 CCAGGAATACATGATCCATTCATGAAAGACAGGGAAGCCCTGGCTCCCTAAGATTCATC 1566  
Dh 1600 CCAGGAATACATGATCCATTCATGAAAGACAGGGAAGCCCTGGCTCCCTAAGATTCATC 1659

Qy 1567 TCTCTCGAACCAAAATTTCTTTCATCTTACAGATGACCCCTCAATCCCATCATPATG 1626  
Dh 1660 TCTCTCGAACCAAAATTTCTTTCATCTTACAGATGACCCCTCAATCCCATCATPATG 1719  
Qy 1627 GTGGGTCCAGGAACCGGCAATGCGGCTTATTTGGGTCTCTTACATATGAGGAATCTC 1686  
Dh 1720 GTGGGTCCAGGAACCGGCAATGCGGCTTATTTGGGTCTCTTACATATGAGGAATCTC 1779  
Qy 1687 CAGAACCAACCCAGATGGAATTTTGGAGCAATGTGTGTTTTTGGCTGCAAGCAT 1746  
Dh 1780 CAGAACCAACCCAGATGGAATTTTGGAGCAATGTGTGTTTTTGGCTGCAAGCAT 1839  
Qy 1747 AAGGATAGGATTTATTTATTTCAAGAAAAGCTCAGCATTTTCTTAAAGCATGGGATCTTA 1806  
Dh 1840 AAGGATAGGATTTATTTATTTCAAGAAAAGCTCAGCATTTTCTTAAAGCATGGGATCTTA 1899  
Qy 1807 ACTCATCTAAAGGTTCTTCTCAAGAGATGCTCTGTTGGGGAGGAGAGAGCCCAAGCA 1866  
Dh 1900 ACTCATCTAAAGGTTCTTCTCAAGAGATGCTCTGTTGGGGAGGAGAGAGCCCAAGCA 1959  
Qy 1867 AAGTATGTACAGACAAATCCAGCTTCAATGGCCAGACAGGTGGCGAATCTCTCCAG 1926  
Dh 1960 AAGTATGTACAGACAAATCCAGCTTCAATGGCCAGACAGGTGGCGAATCTCTCCAG 2019  
Qy 1927 GAGAACGCGCATATTTATGTGTGTGAGATGCAAGAAATATGGCCAAAGATGTACATGAT 1986  
Dh 2020 GAGAACGCGCATATTTATGTGTGTGAGATGCAAGAAATATGGCCAAAGATGTACATGAT 2079  
Qy 1987 GCCCTGTGCAAAATATATAGCAAAAGAGTTGAGTTGAAAACCTAAGCAATGAAAACC 2046  
Dh 2080 GCCCTGTGCAAAATATATAGCAAAAGAGTTGAGTTGAAAACCTAAGCAATGAAAACC 2139  
Qy 2047 CTGGCCACTTTTAAAGAAAGAAAACGCTACTCTTCAAGATATTTGTCTATTA 2097  
Dh 2140 CTGGCCACTTTTAAAGAAAGAAAACGCTACTCTTCAAGATATTTGTCTATTA 2190

RESULT 14  
ADQ39030  
ID ADQ39030 standard; DNA; 3274 BP.  
XX  
AC ADQ39030;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Human SNP containing myocardial infarction-associated gene, SEQ ID 693.  
XX  
KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;  
KW cardiant; gene therapy; human; gene; db.  
OS Homo sapiens.  
XX  
PN W02004058052-A2.  
XX  
PD 15-JUL-2004.  
XX  
PF 22-DEC-2003; 2003MO-US040978.  
XX  
PR 20-DEC-2002; 2002US-0434778P.  
PR 10-MAR-2003; 2003US-0453135P.  
PR 30-APR-2003; 2003US-0466412P.  
PR 23-SEP-2003; 2003US-0504955P.  
XX  
PA (APPL-) APPLERA CORP.  
XX  
PI Cargill M, Devlin JI, Iakubova O;  
XX  
XX WPI; 2004-533949/51.  
DR P-PDB; ADQ39858.  
XX  
XX Identifying an individual who has an altered risk for developing  
PT myocardial infarction by detecting a single nucleotide polymorphism in

PT the individual's nucleic acids.

PS Claim 7; SEQ ID NO 693; 145pp; English.

The invention relates to a novel method for identifying an individual who has an altered risk for developing myocardial infarction. The method comprises detecting a single nucleotide polymorphism (SNP) in any one of the nucleotide sequences given in the specification in the individual's nucleic acids, where the presence of the SNP is correlated with an altered risk for myocardial infarction in the individual. The invention further comprises: an isolated nucleic acid molecule comprising at least 8 contiguous nucleotides where one of the nucleotides is an SNP given in the specification or its complement and encoding any one of the amino acid sequences given in the specification; an isolated polypeptide comprising an amino acid sequence given in the specification; an antibody that specifically binds to the polypeptide or its antigen-binding fragment; an amplified polynucleotide containing an SNP given in the specification and which is between about 16 and 1000 nucleotides in length; a kit for detecting an SNP in a nucleic acid, comprising the polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a nucleic acid molecule; a method of detecting a variant polypeptide; and a method for identifying an agent useful in treating or preventing myocardial infarction. The novel detection method has cardiant activity. The nucleic acids of the invention may be used in gene therapy. The method is useful in identifying an individual who has an increased or decreased risk for developing myocardial infarction and for preparing a composition for treating or preventing myocardial infarction. This polynucleotide sequence represents a human myocardial infarction-associated gene containing one or more SNPs of the invention. Note: This sequence was not shown in the specification. The sequence has come from an electronic sequence listing downloaded from the WIPO website.

Sequence 3274 BP; 932 A; 694 C; 672 G; 946 T; 0 U; 30 Other;

Query Match	50.64;	Score 1062;	DB 13;	Length 3274;
-------------	--------	-------------	--------	--------------

Matches 2012; Conservative 0; Mismatches

Matches 2012; Conservative	0;	Mismatches	19;	Indels	0;	Gaps	0;
----------------------------	----	------------	-----	--------	----	------	----

QY	67	TGTGAGCAAGCTGTGGTACATGGAATTTTCTGCAGATCTTCACTGATTAATGAATCCGAT	126
Db	178	TGTGAGCAAGCTGTGGTACATGGAATTTTCTGCAGATCTTCACTGATTAATGAATCCGAT	237
QY	127	AAGTATGACCTTAAAAACCGAAACACTCTCTCTTGTGTGTGTGGTTTCTACACGGGACAC	186
Db	238	AAGTATGACCTTAAAAACCGAAACACTCTCTCTTGTGTGTGTGGTTTCTACACGGGACAC	297
QY	187	GGAGACCAACCCGACACAGACCCCGCAAGTTTGTATAAGAAATACAGAACCAACCTGCGCG	246
Db	298	GGAGACCAACCCGACACAGACCCCGCAAGTTTGTATAAGAAATACAGAACCAACCTGCGCG	357
QY	247	GTTGATTTCTTTTGTCTCACTCGCGGTATGGGTTACTGGTCTCGGTGATTCAGATATACCC	306
Db	358	GTTGATTTCTTTTGTCTCACTCGCGGTATGGGTTACTGGTCTGTGTATTCAGATATACACC	417
QY	307	TACTTTTGCATGTGGGGGGAAGATTAATTGATTAACGACTTCAAGAGCTTGGAGCCGGCAAT	366
Db	418	TACTTTTGCATGTGGGGGGAAGATTAATTGATTAAYGACTTCAAGAGCTTGGAGCCGGCAAT	477
QY	367	TTCTATGACATGCGACATGACAGATGACTGTGTATAGTTTGAACCTGTGTGTGACCCGTGG	426
Db	478	TTCTATGACATGCGACATTCACATGATGACTGTGTATAGTTTGAACCTGTGTGTGACCCGTGG	537
QY	427	ATTGCTGACACTGTGGCCAGCCCTCAGAAAGCAATTTTAGTCAAGCAGAGAGACAGAGAG	486
Db	538	ATTGCTGACACTGTGGCCAGCCCTCAGAAAGCAATTTTAGTCAAGCAGAGAGACAGAGAG	597
QY	487	ATTAAATGGCGCACTCCCGGTGGCATTCACCTGCATCTTTAGAGACAGACCTTGTGAAGTCA	546
Db	598	ATTAAATGGCGCACTCCCGGTGGCATTCACCTGCATCTTGAAGACAGACCTTGTGAAGTCA	657
QY	547	GAGCTGTCTACACATTGAATCTCAAGTCGAGCTTCTGAGATTTGATGATTCAGAGAAAG	606
Db	658	GAGCTGTCTACACATTGAATCTCAAGTCGAGCTTCTGAGATTTGATGATTCAGAGAAAG	717

QY	607	GATTCGAGGTTTGGAGCAAAATGCAGTAAACGAAACCAATCCANTGTTGTAATGAA	666
Db	718	GATTCTAGGTTTGAAGCAAAATGCAGTAAACGAAACCAATCCANTGTTGTAATGAA	777
QY	667	GACTTGAAGTCTCACTTACCCGTTCCGATACCCCACTCTCAACAAGCTCTGTAAATAT	726
Db	778	GACTTGAAGTCTCACTTACCCGTTCCGATACCCCACTCTCAACAAGCTCTGTAAATAT	837
QY	727	CTGTGTTTACCCCGAATATTTACAGGTACATGTGACAGAGTCTTTGGCCAGAGGAA	786
Db	838	CTGTGTTTACCCCGAATATTTACAGGTACATGTGACAGAGGTCCTTGGCCAGAGGAA	897
QY	787	AGCGAAGTAATGTGACTTCAGCAGATCCAGTTTTCAGAGGCCAATTTTCAAGACAGTT	846
Db	898	AGCGAAGTAATGTGACTTCAGCAGATCCAGTTTTCAGAGGCCAATTTTCAAGAGCAGTT	957
QY	847	CAACTTACTACGAATGATGCCATTAATAACACTGTGCTGATGAATTTGACATTTTCAAT	906
Db	958	CAACTTACTACGAATGATGCCATTAATAACACTGTGCTGATGAATTTGACATTTTCAAT	1017
QY	907	ACAACTTTTCTTATCAGCCTGGAGATGCTTCAGGCGTATCTGSCCTTACAGATATCT	966
Db	1018	ACAACTTTTCTTATCAGCCTGGAGATGCTTCAGGCGTATCTGSCCTTACAGATATCT	1077
QY	967	GAGGTACAAAGCTACTCTCAAAAGACTCAGCTTGAATTAATAAAGACACTGCGTCTT	1026
Db	1078	GAGGTACAAAGCTACTCTCAAAAGACTCAGCTTGAATTAATAAAGACACTGCGTCTT	1137
QY	1027	TTGAAATTAAGGCAACACAAAGAAAGAGACTTACCCCAAGCATATACCTGCG	1086
Db	1138	TTGAAATTAAGGCAACACARBAAGAAAGAGACTTACCCCAAGCATATACCTGCG	1197
QY	1087	GGATGTTCTCCCAAGTTCAATTTTACCGAGTCTTGAATTCGAGCAATCTTAAAG	1146
Db	1198	GGATGTTCTCCCAAGTTCAATTTTACCGAGTCTTGAATTCGAGCAATCTTAAAG	1257
QY	1147	GCATTTTTCAGAGCCTTGTGACTATACCAAGTACAGTGTGTAAAGCCGAGGCTACAG	1206
Db	1258	GCATTTTTCAGAGCCTTGTGACTATACCAAGTACAGTGTGTAAAGCCGAGGCTACAG	1317
QY	1207	GAGCTGTGCAATTAACAAGGGGACGCCGATTATACCGCCTTTGTATCCAGATGCCCTGACC	1266
Db	1318	GAGCTGTGCAATTAACAAGGGGACGCCGATTATACCGCCTTTGTATCCAGATGCCCTGACC	1377
QY	1267	TGCTTGTGATATCTCTCTCGCTTCCCTTCTTGACAGCCACACTCAGTCTCTGCTC	1326
Db	1378	TGCTTGTGATATCTCTCTCGCTTCCCTTCTTGACAGCCACACTCAGTCTCTGCTC	1437
QY	1327	GAACATCTTCTTAAACTTCAACCCAGACCAATATTCGTGTGCAAGTCAAGTTTATTTAC	1386
Db	1438	GAACATCTTCTTAAACTTCAACCCAGACCAATATTCGTGTGCAAGTCAAGTTTATTTAC	1497
QY	1387	CCAGGAAGCTCCCAATTTTGTCTTCAAACTGTGAAATTTCTGTACTGACCAACAAG	1446
Db	1498	CCAGGAAGCTCCCAATTTTGTCTTCAAACTGTGAAATTTCTGTACTGACCAACAAG	1557
QY	1447	GTTCTGCGGAAGGAGATATGTACAGGCTGAGCTGAGCTTGTGTGTTCTTCAAGTTCTTCA	1506
Db	1558	GTTCTGCGGAAGGAGATATGTACAGGCTGAGCTGAGCTTGTGTGTTCTTCAAGTTCTTCA	1617
QY	1507	CCAAACATATCATGTATCCATGAAGACAGGCGGAAAGCCCGGCTCCCTAATATTCATC	1566
Db	1618	CCAAACATATCATGTATCCATGAAGACAGGCGGAAAGCCCGGCTCCCTAATATTCATC	1677
QY	1567	TCTCCCTGAAACAATAATCTTCTTCACTTACAGATGACCCCTCATCTCCCATATATG	1626
Db	1678	TCTCCCTGAAACAATAATCTTCTTCACTTACAGATGACCCCTCATCTCCCATATATG	1737
QY	1627	GTCGGTCCAGAACCCGCGATAGCCCGTTTATTTGGTCTCTACCAATATAGAGAAATCTC	1686
Db	1738	GTCGGTCCAGAACCCGCGATAGCCCGTTTATTTGGTCTCTACCAATATAGAGAAATCTC	1797



```
Db      832  ||||| 891
      832  AAATTCCTGGTTTACCCCGAATATTTCAGGTACATCTGCAGGAGTCTTTGGCCAG
Qy      781  GAGGAAAGCCAGTATCTGTGACTTCAGCAGATCCAGTTTTCAGTGCATTTCAAG 840
      781  |||||
Db      892  GAGGAAAGCCAGTATCTGTGACTTCAGCAGATCCAGTTTTCAGTGCATTTCAAG 951
      892  |||||
Qy      841  GCAGTTCACCTTCTAGATGATGCGCATPAAAAACCACTCTGCTGGTAGAATTGACATT 900
      841  |||||
Db      952  GCAGTTCACCTTCTAGATGATGCGCATPAAAAACCACTCTGCTGGTAGAATTGACATT 1011
      952  |||||
Qy      901  TCAATATACAGACTTTTCTATCAGCCTGAGATGCCCTTCAGCGTGAATGCTCCTAACAGT 960
      901  |||||
Db      1012 TCATATACAGACTTTTCTATCAGCCTGAGATGCCCTTCAGCGTGAATGCTCCTAACAGT 1071
      1012 |||||
Qy      961  GATTCTGAGGTACAAAGCCTACTCCAAAGACTGCAAGCTTGAAGATPAAAAGAGCACTGC 1020
      961  |||||
Db      1072 GATTCTGAGGTACAAAGCCTACTCCAAAGACTGCAAGCTTGAAGATPAAAAGAGCACTGC 1131
      1072 |||||
Qy      1021 GTCCCTTTGAAAATAAAGCGACACAAAGAGAAAGG 1058
      1021 |||||
Db      1132 GTCCCTTTGAAAATAAAGCGACACAAAGAGAAAGG 1169
      1132 |||||
```

Search completed: August 27, 2005, 01:18:31  
Job time : 738.49 secs

**This Page Blank (usp10)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: August 26, 2005, 23:30:25 ; Search time 235.757 Seconds  
(without alignments)  
14554.251 Million cell updates/sec

Title: US-09-371-347A-41

Perfect score: 2097  
Sequence: 1 atcgagagaggtctctgtact.....ttcagatattcgtcataa 2097

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA:  
1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/prodata/1/ina/6C\_COMB.seq:\*  
6: /cgn2\_6/prodata/1/ina/6D\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2046	97.6	3259	3 US-09-318-448-23	Sequence 23, Appl
2	1995	95.1	3242	4 US-09-949-016-4215	Sequence 4215, Ap
3	386	18.4	390	3 US-08-905-223-71	Sequence 71, Appl
4	330	15.7	601	4 US-09-949-016-150019	Sequence 150019, A
5	330	15.7	35916	4 US-09-949-016-15957	Sequence 15957, A
6	279	13.3	601	4 US-09-949-016-150020	Sequence 150020, A
7	189	9.0	601	4 US-09-949-016-150037	Sequence 150037, A
8	158	7.5	2475	4 US-09-566-921-88	Sequence 88, Appl
9	155	7.4	601	4 US-09-949-016-150030	Sequence 150030, A
10	145	6.9	601	4 US-09-949-016-150031	Sequence 150031, A
11	137	6.5	601	4 US-09-949-016-150046	Sequence 150046, A
12	137	6.5	601	4 US-09-949-016-150047	Sequence 150047, A
13	125	6.0	601	4 US-09-949-016-150029	Sequence 150029, A
14	121	5.8	601	4 US-09-949-016-150041	Sequence 150041, A
15	121	5.8	601	4 US-09-949-016-150042	Sequence 150042, A
16	119	5.7	601	4 US-09-949-016-150008	Sequence 150008, A
17	119	5.7	601	4 US-09-949-016-150055	Sequence 150055, A
18	110	5.2	601	4 US-09-949-016-150048	Sequence 150048, A
19	94	4.5	601	4 US-09-949-016-150032	Sequence 150032, A
20	78	3.7	601	4 US-09-949-016-150007	Sequence 150007, A
21	75	3.6	601	4 US-09-949-016-150018	Sequence 150018, A
22	65	3.1	244	4 US-09-471-276-495	Sequence 495, App
23	30	1.4	1681	4 US-09-023-655-453	Sequence 453, App
24	20	1.0	273	4 US-09-513-999C-14761	Sequence 14761, A
25	20	1.0	440	4 US-09-397-787-305	Sequence 305, App
26	20	1.0	444	4 US-09-621-976-14139	Sequence 14139, A
27	20	1.0	445	3 US-09-397-787-274	Sequence 274, App

C	28	20	1.0	174259	4 US-09-949-016-11968	Sequence 11968, A
C	29	20	1.0	174262	4 US-09-949-016-14259	Sequence 14259, A
C	30	19	0.9	169	1 US-08-166-346A-8	Sequence 8, Appl
C	31	19	0.9	459	4 US-09-621-976-8324	Sequence 8324, Ap
C	32	19	0.9	3969	3 US-09-518-386B-4	Sequence 4, Appl
C	33	19	0.9	4396	3 US-09-821-736-1	Sequence 1, Appl
C	34	19	0.9	14721	4 US-09-949-016-13507	Sequence 13507, A
C	35	19	0.9	25199	4 US-09-949-016-13361	Sequence 13361, A
C	36	19	0.9	129658	4 US-09-949-016-17195	Sequence 17195, A
C	37	19	0.9	186734	4 US-09-949-016-14870	Sequence 14870, A
C	38	19	0.9	193689	4 US-09-949-016-12350	Sequence 12350, A
C	39	19	0.9	193689	4 US-09-949-016-13088	Sequence 13088, A
C	40	19	0.9	200663	4 US-09-949-016-12569	Sequence 12569, A
C	41	19	0.9	203093	4 US-09-949-016-14445	Sequence 14445, A
C	42	18	0.9	78	2 US-08-749-852-56	Sequence 56, Appl
C	43	18	0.9	78	2 US-08-749-852-58	Sequence 58, Appl
C	44	18	0.9	511	4 US-09-902-540-1374	Sequence 1374, Ap
C	45	18	0.9	531	4 US-09-252-991A-2223	Sequence 2223, Ap

#### ALIGNMENTS

RESULT 1						
US-09-318-448-23						
Sequence 23, Application US/09318448						
Patent No. 6210950						
GENERAL INFORMATION:						
APPLICANT: Johnson, William G.						
TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING						
TITLE OF INVENTION: DEVELOPMENTAL DISORDERS						
FILE REFERENCE: 601-1-057						
CURRENT APPLICATION NUMBER: US/09/318,448						
CURRENT FILING DATE: 1999-05-25						
NUMBER OF SEQ ID NOS: 46						
SOFTWARE: PatentIn Ver. 2.0						
SEQ ID NO 23						
LENGTH: 3259						
TYPE: DNA						
ORGANISM: Homo sapiens						
US-09-318-448-23						
Query Match						
Best Local Similarity 97.6%; Score 2046; DB 3; Length 3259;						
Matches 2096; Conservative 0; Mismatches 1; Indels 0; Gaps 0;						
Qy	1	ATGAGAGGTTTCTGTTACTATGCTACACAGCAGCAGCAAGCCATCGCAGAA	60			
Db	80	ATGAGAGGTTTCTGTTACTATGCTACACAGCAGCAGCAAGCCATCGCAGAA	139			
Qy	61	GAATATGTGAGCAAGCTGTGTACATGATTTTTCGAGATCTTCACTGTATTAGTAA	120			
Db	140	GAATATGTGAGCAAGCTGTGTACATGATTTTTCGAGATCTTCACTGTATTAGTAA	199			
Qy	121	TCCGATTAAGTAACTTAACCAACCAAGCTCTTGTGTGTTTACACAG	180			
Db	200	TCCGATTAAGTAACTTAACCAACCAAGCTCTTGTGTGTTTACACAG	259			
Qy	181	GGCAGCGAGACCCACCCGACAGCAGCCGAGTTTGAAGAAATACAGAACAAACA	240			
Db	260	GGCAGCGAGACCCACCCGACAGCAGCCGAGTTTGAAGAAATACAGAACAAACA	319			
Qy	241	CTCCGCGTGAATTTCTTGTCTACCTCCGATAGGTTTACTGGCTTCGGTATTCAGAA	300			
Db	320	CTCCGCGTGAATTTCTTGTCTACCTCCGATAGGTTTACTGGCTTCGGTATTCAGAA	379			
Qy	301	TAACTACTTTTTCGATGGGCGAGATTAATGAATTAAGAGCTTGAGGCC	360			
Db	380	TAACTACTTTTTCGATGGGCGAGATTAATGAATTAAGAGCTTGAGGCC	439			
Qy	361	CGCATTTTCATGACACTGACATGACATGATGATGATTTTAAACTTGTGTTGAG	420			

Dh 440 CCGCATTTCATGACACTGACATGACATGACTGTAGTTTGAACCTTGCTTGAG 439  
Qy 421 CCGTGATGCTGACACTGTGCGCAGCCCTGAGAAAGCATTTTAGTTCAGACAGAGCA 480  
Dh 500 CCGTGATGCTGACACTGTGCGCAGCCCTGAGAAAGCATTTTAGTTCAGACAGAGCA 559  
Qy 481 GAGAGATTAAGTGGCGCACTCCCGGTGGCATCACTGCACTCTTGAGAGACAGACTTGTG 540  
Dh 560 GAGAGATTAAGTGGCGCACTCCCGGTGGCATCACTGCACTCTTGAGAGACAGACTTGTG 619  
Qy 541 AAGTCAGAGCTGTCACATTTGAATCTCAAGTCAGCTTCGAGATTCGATTCAGAG 600  
Dh 620 AAGTCAGAGCTGTCACATTTGAATCTCAAGTCAGCTTCGAGATTCGATTCAGAG 679  
Qy 601 AGAAAGATTCGAGCTTTGAGCAAAATGCAAGTGAACAGCAACCAATCCATGTTGA 660  
Dh 680 AGAAAGATTCGAGCTTTGAGCAAAATGCAAGTGAACAGCAACCAATCCATGTTGA 739  
Qy 661 ATTGAAGATTTGAGTCTCTCACTTACCCGTTCCGTACCCCACTTCACAGGCTCTCTG 720  
Dh 740 ATTGAAGATTTGAGTCTCTCACTTACCCGTTCCGTACCCCACTTCACAGGCTCTCTG 799  
Qy 721 AATATCTGCTGATTAACCCCGAGATATTTACAGGTATGCTGACAGAGTCTCTTGCGCAG 780  
Dh 800 AATATCTGCTGATTAACCCCGAGATATTTACAGGTATGCTGACAGAGTCTCTTGCGCAG 859  
Qy 781 GAGGAAAGCCAAAGTATGCTGACTTCAAGCATCAAGTCTTCAAGTCCAAATTTCAAG 840  
Dh 860 GAGGAAAGCCAAAGTATGCTGACTTCAAGCATCAAGTCTTCAAGTCCAAATTTCAAG 919  
Qy 841 GCAGTTCACTTATAGATGATGCAATTAATAAACACTGCTGCTGATGATGATGATGAT 900  
Dh 920 GCAGTTCACTTATAGATGATGCAATTAATAAACACTGCTGCTGATGATGATGATGAT 979  
Qy 901 TCAAAATACAGACTTTTCTATACAGCTGAGATGCTTCAAGCTGATGCTGCTTCAAGT 960  
Dh 980 TCAAAATACAGACTTTTCTATACAGCTGAGATGCTTCAAGCTGATGCTGCTTCAAGT 1039  
Qy 961 GATTTGAGGTACAAAGCTTACTTCAAAAGCTGAGCTTGAAGTAAAGAGAGCACTGC 1020  
Dh 1040 GATTTGAGGTACAAAGCTTACTTCAAAAGCTGAGCTTGAAGTAAAGAGAGCACTGC 1099  
Qy 1021 GTCTTTTGAATAAAGGAGAGACAAAGAAAGAAAGAGTACCTTACCCAGCATATA 1080  
Dh 1100 GTCTTTTGAATAAAGGAGAGACAAAGAAAGAAAGAGTACCTTACCCAGCATATA 1159  
Qy 1081 CTTGCGGAGATGTTCTCTCAAGTTCATTTTACCTGATGCTTGAATTCGAGCAATTCCT 1140  
Dh 1160 CTTGCGGAGATGTTCTCTCAAGTTCATTTTACCTGATGCTTGAATTCGAGCAATTCCT 1219  
Qy 1141 AAAAAGCATTTTGGAGGCTTGTGACTATACAGTGAACAGTGTGTAAGAAAGCGCAGG 1200  
Dh 1220 AAAAAGCATTTTGGAGGCTTGTGACTATACAGTGAACAGTGTGTAAGAAAGCGCAGG 1279  
Qy 1201 CTACAGAGCTGTGCACTTAAACAGAGGGGACGCAATTAATAGCCGCTTTGTAACAGAGGCC 1260  
Dh 1280 CTACAGAGCTGTGCACTTAAACAGAGGGGACGCAATTAATAGCCGCTTTGTAACAGAGGCC 1339  
Qy 1261 TGTGCTGCTGTTGATGATCTCTCTGCTGCTTCCCTTCTTGCGCAGCACACTGATCTC 1320  
Dh 1340 TGTGCTGCTGTTGATGATCTCTCTGCTGCTTCCCTTCTTGCGCAGCACACTGATCTC 1399  
Qy 1321 CTGCTGCAACATCTTCTTAAACCTTCAACCCAGACCATATTCGTGTGCAAGTCAAGTTTA 1380  
Dh 1400 CTGCTGCAACATCTTCTTAAACCTTCAACCCAGACCATATTCGTGTGCAAGTCAAGTTTA 1459  
Qy 1381 TTTTCAACCAAGAAAGCTCCATTTTGTCTTCAACATTTGTGAAATTTCTGTCTACGCA 1440  
Dh 1460 TTTTCAACCAAGAAAGCTCCATTTTGTCTTCAACATTTGTGAAATTTCTGTCTACGCA 1519  
Qy 1441 ACAGAGGTTCTGCGGAAGGAGTATGTAACAGCTGCTGCGCTTGTGTTGCTTCAAGT 1500  
Dh 1520 ACAGAGGTTCTGCGGAAGGAGTATGTAACAGCTGCTGCGCTTGTGTTGCTTCAAGT 1579

Qy 1501 CTTGACCCAAACATACATGATCCCATGAAAGACGCGGAAAGCCCTGCTCTTAAGATA 1560  
Dh 1580 CTTGACCCAAACATACATGATCCCATGAAAGACGCGGAAAGCCCTGCTCTTAAGATA 1639  
Qy 1561 TCCATCTCTCTGCAACAAATTCCTTCCATTAACAGATGACCCCTCAATCCCATC 1620  
Dh 1640 TCCATCTCTCTGCAACAAATTCCTTCCATTAACAGATGACCCCTCAATCCCATC 1699  
Qy 1621 ATATAGTGGGTCCAGAAACCGGATAGCCCTTTATTTGGTTCCTACAAATAGAGAG 1680  
Dh 1700 ATATAGTGGGTCCAGAAACCGGATAGCCCTTTATTTGGTTCCTACAAATAGAGAG 1759  
Qy 1681 AAATCCCAAGAACACACCCAGATGGAATTTTGAAGCAATGCTGTTTGGCTGC 1740  
Dh 1760 AAATCCCAAGAACACACCCAGATGGAATTTTGAAGCAATGCTGTTTGGCTGC 1819  
Qy 1741 AGGCATTAAGATGAGGATTAATCATTCAGAAAGACCTCAGACATTTCTTAAGATGG 1800  
Dh 1820 AGGCATTAAGATGAGGATTAATCATTCAGAAAGACCTCAGACATTTCTTAAGATGG 1879  
Qy 1801 ATCTTAATCATTAAGGATTTCTTCTCAAGAGATGCTCTGTTGGGAGAGAGAGCC 1860  
Dh 1880 ATCTTAATCATTAAGGATTTCTTCTCAAGAGATGCTCTGTTGGGAGAGAGAGCC 1939  
Qy 1861 CCAGCAAGATATGACAGCAACATCCAGCTTCATGCGCAGAGTGGCGAGATCTTC 1920  
Dh 1940 CCAGCAAGATATGACAGCAACATCCAGCTTCATGCGCAGAGTGGCGAGATCTTC 1999  
Qy 1921 CTCAGAGAAAGGCGCATATTTATGTTGTGAGATGCAAAAGATATGCGCAAGATGTA 1980  
Dh 2000 CTCAGAGAAAGGCGCATATTTATGTTGTGAGATGCAAAAGATATGCGCAAGATGTA 2059  
Qy 1981 CATGATGCCCTTGTGCAAAATATATAGCAAGAGTGTGAGTTGAAAAAATAGAGCATG 2040  
Dh 2060 CATGATGCCCTTGTGCAAAATATATAGCAAGAGTGTGAGTTGAAAAAATAGAGCATG 2119  
Qy 2041 AAAACCTGCGCATTAAAGAAAGAAAGCGTACCTTCAGATATTTGCTCATTA 2097  
Dh 2120 AAAACCTGCGCATTAAAGAAAGAAAGCGTACCTTCAGATATTTGCTCATTA 2176

RESULT 2  
US-09-949-016-4215  
; Sequence 4215, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4215  
; LENGTH: 3242  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-4215

Query Match 95.1%; Score 1995; DB 4; Length 3242;  
Best local Similarity 99.9%; Pred. No. 0;  
Matches 2095; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGAGAGGTTTCTGTTACTATATGCTACACAGAGGACAGCAAGCCATCGCAGAA 60  
|||||

80 ATGAGGAGGTTCTGTTACTATATGCTACAGACAGGAGCAGGAAAGCCATCGAGAA 139  
61 GAAATATGAGAGAGCTGTGTACATGATTTTCTGAGATCTTGACTGTATAGTAA 120  
140 GAAATATGAGAGAGCTGTGTACATGATTTTCTGAGATCTTGACTGTATAGTAA 139  
121 TCGATATGATATGACCTTAAACCCGAAACAGCTCTCTGTTGTGTGTGTTCTACACG 180  
200 TCGATATGATATGACCTTAAACCCGAAACAGCTCTCTGTTGTGTGTGTTCTACACG 259  
181 GGACCGGAGAGCCACCCGACACAGCCCGCAAGTTGTTAAGAAATACAAACAAACA 240  
260 GGACCGGAGAGCCACCCGACACAGCCCGCAAGTTGTTAAGAAATACAAACAAACA 319  
241 CTGCGGTTGATTTCTTTGTCTGACCTGCGGTATAGGTTACTGGTCTCGGTGATTCAGAA 300  
320 CTGCGGTTGATTTCTTTGTCTGACCTGCGGTATAGGTTACTGGTCTCGGTGATTCAGAA 379  
301 TACACCTACTTTTGCATGAGGAGGAGATTAATTGATTAACGACTTCAGAGCTTGAGCC 360  
380 TACACCTACTTTTGCATGAGGAGGAGATTAATTGATTAACGACTTCAGAGCTTGAGCC 439  
361 CGGCAATTTCTATGACACTGACATGCAATGATGATGATGATTTAAGCTTGTGTGAG 420  
440 CGGCAATTTCTATGACACTGACATGCAATGATGATGATGATTTAAGCTTGTGTGAG 499  
421 CGGTGATTTGCTGAGCTGTGCGGACCTTCAGAAAGCTTTTGTGTTCAAGCAAGACA 480  
500 CGGTGATTTGCTGAGCTGTGCGGACCTTCAGAAAGCTTTTGTGTTCAAGCAAGACA 559  
481 GAGGAGATTAAGTGGCCACTCCCGTGGCATCACTGCACTCTTGTGAGACAGACTTGTG 540  
560 GAGGAGATTAAGTGGCCACTCCCGTGGCATCACTGCACTCTTGTGAGACAGACTTGTG 619  
541 AAGTCAGAGCTGTGACATGATTAATCTCAAGTCAGCTTCTGAGATTCAGATTCAGAA 600  
620 AAGTCAGAGCTGTGACATGATTAATCTCAAGTCAGCTTCTGAGATTCAGATTCAGAA 679  
601 AGAAAGATTTCTGAGTTTGAAGCAAAATGCACTGACAGCAACCAATTCATTTGTTGA 660  
680 AGAAAGATTTCTGAGTTTGAAGCAAAATGCACTGACAGCAACCAATTCATTTGTTGA 739  
661 ATTGAAGCTTTGAGCTCTCACTTACCCGTTGGGTACCCCACTCTCAAGAGCTCTG 720  
740 ATTGAAGCTTTGAGCTCTCACTTACCCGTTGGGTACCCCACTCTCAAGAGCTCTG 799  
721 AATATTCCTGTGTTTACCCCGCAATATTTTACAGATCATCTGAGAGATCTCTTGGCAG 780  
800 AATATTCCTGTGTTTACCCCGCAATATTTTACAGATCATCTGAGAGATCTCTTGGCAG 859  
781 GAGGAAAGCCAAATATCTGTGACTTCAGAGATCTCAAGTTTCAAGTCCCAATTTCAAAG 840  
860 GAGGAAAGCCAAATATCTGTGACTTCAGAGATCTCAAGTTTCAAGTCCCAATTTCAAAG 919  
841 GCAGTCAACTTACTGCAATGATGCAATTAACCACTGCTGCTGATTAAGTTGAGCAAT 900  
920 GCAGTCAACTTACTGCAATGATGCAATTAACCACTGCTGCTGATTAAGTTGAGCAAT 979  
901 TCAATATACAGACTTTTCTATACAGCTGAGATGCTTCAGCGTGAATCTGCCCTTAACAGT 960  
980 TCAATATACAGACTTTTCTATACAGCTGAGATGCTTCAGCGTGAATCTGCCCTTAACAGT 1039  
961 GATTCGAGGTACAAAGCTCTTCCAAAGCTGCAAGTTGAAGATTAAGAGAGCACTGC 1020  
1040 GATTCGAGGTACAAAGCTCTTCCAAAGCTGCAAGTTGAAGATTAAGAGAGCACTGC 1099  
1021 GTCTTTTGAATTAAGGACGACACAAAGAAAGAGACTACTTACCCAGCATATA 1080  
1100 GTCTTTTGAATTAAGGACGACACAAAGAAAGAGACTACTTACCCAGCATATA 1159  
1081 CTGCGGAGATTTCTCTCAGATTCATTTTACTGTGTCTTGAATTCGAGCAATTTCT 1140  
1160 CTGCGGAGATTTCTCTCAGATTCATTTTACTGTGTCTTGAATTCGAGCAATTTCT 1219

1141 AAAAAGCATTTTGTGAGCCCTTGTGACTATACAGTGAAGTGTGAAAAAGCCAGG 1200  
1220 AAAAAGCATTTTGTGAGCCCTTGTGACTATACAGTGAAGTGTGAAAAAGCCAGG 1279  
1201 CTACAGAGCTGTGACATTAACAAAGGAGGAGCCGATTAATAGCCGTTTGTACAGATGCC 1260  
1280 CTACAGAGCTGTGACATTAACAAAGGAGGAGCCGATTAATAGCCGTTTGTACAGATGCC 1339  
1261 TGTGCTGCTTGTGATCTCTCTCCGCTTCCCTTCTTGGCAGGCAACACTGATCTC 1320  
1340 TGTGCTGCTTGTGATCTCTCTCCGCTTCCCTTCTTGGCAGGCAACACTGATCTC 1399  
1321 CTGCTGAACATCTTCTTAACTTCAACCCAGACATATTCGTGTCAAGCTCAAGTTTA 1380  
1400 CTGCTGAACATCTTCTTAACTTCAACCCAGACATATTCGTGTCAAGCTCAAGTTTA 1459  
1381 TTTCAACCCAGAAAGCTCATTTTGTCTTCAACATTTGTGAAATTTCTGTCTACTGCAACA 1440  
1460 TTTCAACCCAGAAAGCTCATTTTGTCTTCAACATTTGTGAAATTTCTGTCTACTGCAACA 1519  
1441 ACAGAGTTCTGCGGAGGAGATTAAGTACAGCTGCTGCTTGTGTGTGTTGTTAGTT 1500  
1520 ACAGAGTTCTGCGGAGGAGATTAAGTACAGCTGCTGCTTGTGTGTGTTGTTAGTT 1579  
1501 CTTCAAGCAACATACATGATCCCATGAAAGACAGGAGGAAAGCCCTGCTCTTAAGATA 1560  
1580 CTTCAAGCAACATACATGATCCCATGAAAGACAGGAGGAAAGCCCTGCTCTTAAGATA 1639  
1561 TCCATCTCTCTCTGAAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
1640 TCCATCTCTCTCTGAAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1699  
1621 ATATATGTTGGTCTCAGAAACCCGATAGCCCGTTTATTTGGTCTTACAACTAAGAG 1680  
1700 ATATATGTTGGTCTCAGAAACCCGATAGCCCGTTTATTTGGTCTTACAACTAAGAG 1759  
1681 AAATCTCAAGAAACAAACCCAGATGGAATTTTGGAGCAATGTTGTTTGTGCTGC 1740  
1760 AAATCTCAAGAAACAAACCCAGATGGAATTTTGGAGCAATGTTGTTTGTGCTGC 1819  
1741 AGGCAATAGATTAAGGATTAATCTATTCAGAAAGAGCTCAGACTTCTTAAAGATGG 1800  
1820 AGGCAATAGATTAAGGATTAATCTATTCAGAAAGAGCTCAGACTTCTTAAAGATGG 1879  
1801 ATCTTAATCTAATTAAGGTTTCTTCTCAAGAGATGCTCTGTGTGGAGAGAGAGCC 1860  
1880 ATCTTAATCTAATTAAGGTTTCTTCTCAAGAGATGCTCTGTGTGGAGAGAGAGCC 1939  
1861 CCAGCAAGATTAATTAACAAACATCCAGCTTCAAGGCAAGGAGTGGGAGAAATCCCTC 1920  
1940 CCAGCAAGATTAATTAACAAACATCCAGCTTCAAGGCAAGGAGTGGGAGAAATCCCTC 1999  
1921 CTCCAGAGAAAGGAGGATTAATTAATGTTGTGTGAGATGCAAGAAATATGCAAGATGA 1980  
2000 CTCCAGAGAAAGGAGGATTAATTAATGTTGTGTGAGATGCAAGAAATATGCAAGATGA 2059  
1981 CATGATGCCCTTGTGCAATTAATTAAGCAAAAGGTTGAGTTGAAAATCTAAGAGCAATG 2040  
2060 CATGATGCCCTTGTGCAATTAATTAAGCAAAAGGTTGAGTTGAAAATCTAAGAGCAATG 2119  
2041 AAAAAGCTGCGCACTTTAAAGAAAGAAAGCTACCTCAGGATTTTGTGATTA 2097  
2120 AAAAAGCTGCGCACTTTAAAGAAAGAAAGCTACCTCAGGATTTTGTGATTA 2176

RESULT 3  
US-08-905-223-71  
; Sequence 71, Application US/08905223  
; Patent No. 6222029  
; GENERAL INFORMATION:  
; APPLICANT: Edwards, Jean-Baptiste D.  
; APPLICANT: Duclert, Aymeric

APPLICANT: Lacroix, Bruno  
TITLE OF INVENTION: 5' ESTs FOR SECRETED PROTEINS  
NUMBER OF SEQUENCES: 503  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe, Martens, Olson & Bear  
STREET: 501 West Broadway  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92101-3505  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Win95  
SOFTWARE: Word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/905,223  
FILING DATE:  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Israel, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER:  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 71:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 390 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: DOUBLE  
TOPOLOGY: LINEAR  
MOLECULE TYPE: CDNA  
ORIGINAL SOURCE:  
ORGANISM: Homo Sapiens  
TISSUE TYPE: Brain  
FEATURE:  
NAME/KEY: s1g\_peptide  
LOCATION: 289..357  
IDENTIFICATION METHOD: Von Heijne matrix  
OTHER INFORMATION: score 6.9  
OTHER INFORMATION: seq SLSLASHSVSC/SN  
US-08-905-223-71

Query Match 18.4%; Score 386; DB 3; Length 390;  
Best Local Similarity 100.0%; Pred. No. 4.9e-188;  
Matches 386; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 970 GTACAAAGCTTCTCCAAAGACTGCACTTGAAGATAAAAGAGCACTGCGCTTTTG 1029  
DB 3 GTACAAAGCTTCTCCAAAGACTGCACTTGAAGATAAAAGAGCACTGCGCTTTTG 62  
QY 1030 AAAATAAAGGAGAGACAAAGAAAGAGCTACTTACCCCGCATATACCTGGAGGA 1089  
DB 63 AAAATAAAGGAGAGACAAAGAAAGAGCTACTTACCCCGCATATACCTGGAGGA 122  
QY 1090 TGTTCCTCAGTTCATTTTACCTGCTGTTGAAATCCGAGCAATTCCTAAAAAGCA 1149  
DB 123 TGTTCCTCAGTTCATTTTACCTGCTGTTGAAATCCGAGCAATTCCTAAAAAGCA 182  
QY 1150 TTTTGGAGAGCCCTTGGAGCTATACAGTGAAGAGAGAGAGAGAGAGAGAGAG 1209  
DB 183 TTTTGGAGAGCCCTTGGAGCTATACAGTGAAGAGAGAGAGAGAGAGAGAGAG 242  
QY 1210 CTGAGCAGTAAACAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1269  
DB 243 CTGAGCAGTAAACAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 302  
QY 1270 TTGTTGATCTGAA 1329  
DB 303 TTGTTGATCTGAA 362  
QY 1330 CATCTTCTAAACTTCAACCCAGACC 1355

DB 363 CATCTTCTAAACTTCAACCCAGACC 388

RESULT 4  
US-09-949-016-150019  
Sequence 150019, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 150019  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-150019

Query Match 15.7%; Score 330; DB 4; Length 601;  
Best Local Similarity 99.7%; Pred. No. 3.3e-159;  
Matches 380; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 401 GTTTAGAACTTGGTGGTGAAGCCGTGATGCTGCACTTGCGCCGCTCAGAAACATT 460  
DB 178 GTTTAGAACTTGGTGGTGAAGCCGTGATGCTGCACTTGCGCCGCTCAGAAACATT 237  
QY 461 TTAGTCAAGCAGAGAGACAAAGAGATAGTGGCGCACTCCGGTGGCATCCTGCAT 520  
DB 238 TTAGTCAAGCAGAGAGACAAAGAGATAGTGGCGCACTCCGGTGGCATCCTGCAT 297  
QY 521 CCTTGAGAGACAGCCTTGGAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 580  
DB 298 CCTTGAGAGACAGCCTTGGAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 357  
QY 581 TGAGATTCAGATTCAGAGAAAGAAAGATTCAGAGTTTGAAGCAAAATGACAGTACA 640  
DB 358 TGAGATTCAGATTCAGAGAAAGAAAGATTCAGAGTTTGAAGCAAAATGACAGTACA 417  
QY 641 GCAACCAATCCAAATGTTGTAATTAAGACTTTGAGTCTCACTTACCCGTTGGTACCC 700  
DB 418 GCAACCAATCCAAATGTTGTAATTAAGACTTTGAGTCTCACTTACCCGTTGGTACCC 477  
QY 701 CACTTCACAAGCCTCTGGAATATTCGCTTACCCCGAATTTTACAGAGTACATC 760  
DB 478 CACTTCACAAGCCTCTGGAATATTCGCTTACCCCGAATTTTACAGAGTACATC 537  
QY 761 TGCAGAGTCTCTTGGCCAGG 781  
DB 538 TGCAGAGTCTCTTGGCCAGG 558

RESULT 5  
US-09-949-016-15957  
Sequence 15957, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 15957  
LENGTH: 35916  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-15957

Query Match 15.7%; Score 330; DB 4; Length 35916;

Best Local Similarity 99.7%; Pred. No. 4.2e-159;  
Matches 380; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 401 GTTTAGAACTTGTGTTGAGCCGTGATGCTGCACTGCGCCAGCCCTCAGAAAGCATT 460  
DB 10781 GTTTAGAACTTGTGTTGAGCCGTGATGCTGCACTGCGCCAGCCCTCAGAAAGCATT 10840  
QY 461 TTAGGTCAAGCAGAGCAAGAGAGATTAAGTGGCCACTCCCGGTGGATCACTGCAT 520  
DB 10841 TTAGGTCAAGCAGAGCAAGAGAGATTAAGTGGCCACTCCCGGTGGATCACTGCAT 10900  
QY 521 CCTTGAGCAGACCTTGTGAAGTCAAGCTGCTACATTTGAATCTCAAGTCGACCTTC 580  
DB 10901 CCTTGAGCAGACCTTGTGAAGTCAAGCTGCTACATTTGAATCTCAAGTCGACCTTC 10960  
QY 581 TGAGATTGATGATTCAGAGAAAGAGATTCTGAGGTTTGAAGCAAAATGACGTAACA 640  
DB 10961 TGAGATTGATGATTCAGAGAAAGAGATTCTGAGGTTTGAAGCAAAATGACGTAACA 11020  
QY 641 GCAACCAATCCAAATGTTGATTAATGAAGTCTCACTTACCCGTTGCGTACCCC 700  
DB 11021 GCAACCAATCCAAATGTTGATTAATGAAGTCTCACTTACCCGTTGCGTACCCC 11080  
QY 701 CACTCTCAAGCCTCTCTGAATATTCCTGTTTACCCCAAAATTTTACAGGTACATC 760  
DB 11081 CACTCTCAAGCCTCTCTGAATATTCCTGTTTACCCCAAAATTTTACAGGTACATC 11140  
QY 761 TGCAGAGTCTCTGGCCAGG 781  
DB 11141 TGCAGAGTCTCTGGCCAGG 11161

## RESULT 6

US-09-949-016-150020  
Sequence 150020, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 150020  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-150020

Query Match 13.3%; Score 279; DB 4; Length 601;

Best Local Similarity 99.5%; Pred. No. 5.9e-133;  
Matches 379; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 401 GTTTAGAACTTGTGTTGAGCCGTGATGCTGCACTGCGCCAGCCCTCAGAAAGCATT 460  
DB 10781 GTTTAGAACTTGTGTTGAGCCGTGATGCTGCACTGCGCCAGCCCTCAGAAAGCATT 10840  
QY 461 TTAGGTCAAGCAGAGCAAGAGAGATTAAGTGGCCACTCCCGGTGGATCACTGCAT 520  
DB 10841 TTAGGTCAAGCAGAGCAAGAGAGATTAAGTGGCCACTCCCGGTGGATCACTGCAT 10900  
QY 521 CCTTGAGCAGACCTTGTGAAGTCAAGCTGCTACATTTGAATCTCAAGTCGACCTTC 580  
DB 10901 CCTTGAGCAGACCTTGTGAAGTCAAGCTGCTACATTTGAATCTCAAGTCGACCTTC 10960  
QY 581 TGAGATTGATGATTCAGAGAAAGAGATTCTGAGGTTTGAAGCAAAATGACGTAACA 640  
DB 10961 TGAGATTGATGATTCAGAGAAAGAGATTCTGAGGTTTGAAGCAAAATGACGTAACA 11020  
QY 641 GCAACCAATCCAAATGTTGATTAATGAAGTCTCACTTACCCGTTGCGTACCCC 700  
DB 11021 GCAACCAATCCAAATGTTGATTAATGAAGTCTCACTTACCCGTTGCGTACCCC 11080  
QY 701 CACTCTCAAGCCTCTCTGAATATTCCTGTTTACCCCAAAATTTTACAGGTACATC 760  
DB 11081 CACTCTCAAGCCTCTCTGAATATTCCTGTTTACCCCAAAATTTTACAGGTACATC 11140  
QY 761 TGCAGAGTCTCTGGCCAGG 781  
DB 11141 TGCAGAGTCTCTGGCCAGG 11161

## RESULT 7

US-09-949-016-150037  
Sequence 150037, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 150037  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-150037

Query Match 9.0%; Score 189; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 1.2e-86;  
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1369 AGCTCAAGTTTATTTTCAACCCAGAAAGCTCTTTTGTCTTCAACATTTGTGAATTTCTG 1428  
DB 18 AGCTCAAGTTTATTTTCAACCCAGAAAGCTCTTTTGTCTTCAACATTTGTGAATTTCTG 77  
QY 1429 TCTACTGCCAACAAGAGGTTCTGCGAAGGAGATATGTACAGGCTGCTGCTGTG 1488  
DB 78 TCTACTGCCAACAAGAGGTTCTGCGAAGGAGATATGTACAGGCTGCTGCTGTG 137  
QY 1489 GTTGCTTCAAGTTCTTCAAGCAACATCATGATCCCATGAAGACAGCGGAAAGCCCTG 1548  
DB 138 GTTGCTTCAAGTTCTTCAAGCAACATCATGATCCCATGAAGACAGCGGAAAGCCCTG 197

QY	1549	GCTCCTAAG	1557
Db	198	GCTCCTAAG	206

RESULT 8  
US-09-566-921-88  
; Sequence 88, Application US/09566921

QY 525 GAGGACAGACCTTGATGAAGTCAGAGCTGCAACAATTGAATCTCAAGTCAGCTCTGAG 584  
 Db 16 GAGGACAGACCTTGATGAAGTCAGAGCTGCAACAATTGAATCTCAAGTCAGCTCTGAG 75

Dy 585 ATTCGATGATT CAGAGAAGAGATTCTGAGTTTGAAGCAAAATGCAGTGAAACAGCAA 644  
Db 76 ATTCGATGATT CAGAGAAGAGATTCTGAGTTTGAAGCAAAATGCAGTGAAACAGCAA 135

Qy 645 CCATCCATGTTGTAATTGAGACTTGTGCTCTAC 682  
|||  
Db 136 CCAATCCATGTTGTAATTGAGACTTGTGCTCTAC 173

```

RESULT 9
US-09-949-016-150030
; Sequence 150030, Application US/09949016
; Patent No. 681239
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 150030
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-150030

```

```

Query Match      7.4%; Score 155; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 3.8e-69;
Matches 155; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY	904	AAATCAGACCTTTTCCCTATCAGCCCTGGAGATCCCTTCACGCGGATCTGCCCTTAACAGTGAT	963
QY	320	AATACAGACCTTTTCCCTATCAGCCCTGGAGATCCCTTCACGCGGATCTGCCCTTAACAGTGAT	379
QY	964	TTCTGAGGTACAAAGCCCTACTCCAAAGACTCGAGCTTGAAGATTAAGAAGAGACACTCGCTC	1022
Db	380	TTCTGAGGTACAAAGCCCTACTCCAAAGACTCGAGCTTGAAGATTAAGAAGAGACACTCGCTC	439
QY	1024	CTTTTGAATAATTAAGGACACACAAGAAGAAGG	1058
Db	440	CTTTTGAATAATTAAGGACACACAAGAAGAAGG	474

```

RESULT 10
US-09-949-016-150031
: Sequence 150031, Application US/09949016
: Patent No. 6812339
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: C0001307
: CURRENT APPLICATION NUMBER: US/09/949,016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 150031
: LENGTH: 601
: TYPE: DNA
: ORGANISM: Human
US-09-949-016-150031

```

Query Match	6.9%	Score 145;	DB 4;	Length 601;
Best Local Similarity	100.0%;	Pred. No. 5.4e-64;		
Matches 145; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0

[illegible][illegible]

```

RESULT 11
US-09-949-016-150046
; Sequence 150046, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

```

```

:
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
:
: SOFTWARE: FastSeq for Windows Version 4.0.
:
: SEQ ID NO: 150046
:
: LENGTH: 601
:
: TYPE: DNA
:
: ORGANISM: Human
:
US-09-949-016-150046

```

Query Match	6.5%	Score 137;	DB 4;	Length 601;
Best Local Similarity	99.5%;	Pred. No. 7e-60;		
Matches 187;	Conservative	0;	Mismatches 1;	Indels 0;
				Gaps 0;

QY	1765	TTGAGAAAAGAGCTCAGACATTTCTTTAAGATGGAGACTTAACTCACTAAAGGTTTCC	1824
Db	413	TTTGAAAAAGAGCTCAGACATTTCTTTAAGATGGAGACTTAACTCACTAAAGGTTTCC	472
QY	1825	TTCTCAAGAGATGCTCTGTTGGGAGAGGAAGCCCCAGCAAACTATGTACAAAGCAAC	1884
Db	473	TTCTCAAGAGATGCTCTGTTGGGAGAGGAAGCCCCAGCAAACTATGTGCAAGCAAC	532
QY	1885	ATCCAGCTTCAATGCCCAGACAGGTGGCCGAATCTCTCTCCAGAGAAAGGCGCATATTTAT	1944
Db	533	ATCCAGCTTCAATGCCCAGACAGGTGGCCGAATCTCTCTCCAGAGAAAGGCGCATATTTAT	592
QY	1945	GTGTGTGG 1952	
Db	593	GTGTGTGG 600	

RESULT 12  
 US-09-949-016-150047  
 : Sequence.150047, Application US/09949016  
 : Patent No.6812339  
 : GENERAL INFORMATION:  
 : APPLICANT: VENTER, J. Craig et al.  
 : TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 : TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 : FILE REFERENCE: CL001307  
 : CURRENT APPLICATION NUMBER: US/09/949,016  
 : CURRENT FILING DATE: 2000-04-14  
 : PRIOR APPLICATION NUMBER: 60/241,755  
 : PRIOR FILING DATE: 2000-10-20  
 : PRIOR APPLICATION NUMBER: 60/237,768  
 : PRIOR FILING DATE: 2000-10-03  
 : PRIOR APPLICATION NUMBER: 60/231,498  
 : PRIOR FILING DATE: 2000-03-08  
 : NUMBER OF SEQ ID NOS: 207012  
 : SOFTWARE: FASTSEQ for Windows Version 4.0  
 : SEQ ID NO 150047  
 : LENGTH: 601  
 : TYPE: DNA  
 : ORGANISM: Human  
 : US-09-949-016-150047

	Query Match	Best Local Similarity	6.5%	Score 137;	DB 4;	Length 601;
	Matches	187;	Conservative	0;	Mismatches	1;
					Indels	0;
					Gaps	0;
Qy	1765	TTGAGAAAAGAGCTCAGACATTTCTTAAAGCATGGATCTTAACATCTAAAGTTTCC	1824			
Db	191	TTGAGAAAAGAGCTCAGACATTTCTTAAAGCATGGATCTTAACATCTAAAGTTTCC	250			
Qy	1825	TTCTCAAGAGATGCTCTCTGTTGGGAGAGGAAAGCCCGACAAAGTATGTACAGACAAAC	1888			
Db	251	TTCTCAAGAGATGCTCTCTGTTGGGAGAGGAAAGCCCGACAAAGTATGTACAGACAAAC	310			
Qy	1885	ATCCAGCTTCATGGCCAGAGAGGTGGCGAAGATCTCTCTCAGAGAAAGGCGCATTTTAT	1944			
Db	311	ATCCAGCTTCATGGCCAGAGAGGTGGCGAAGATCTCTCTCAGAGAAAGGCGCATTTTAT	370			

Qy	1945	GTGTGTGG	1952
Db	371	GTGTGTGG	378

```

RESULT 13
US-09-949-016-150029
/ Sequence 150029, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: C0001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 150029
/ LENGTH: 601
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-150029

```

	Query Match	Similarity	Score	DB	Length
Best Local	100.0%	1.1e-53			
Matches	125	Conservative	0	Mismatches	0
				Indels	0
				Gaps	0

  

	Query	Match	Similarity	Score	DB	Length
779	AGAGGAAAGCCAGATATCTGTGACTTCAGCAGATCCAGTTTTCAGAGTGCCAATTTCAA	838				
379	AGAGGAAAGCCAGATATCTGTGACTTCAGCAGATCCAGTTTTCAGAGTGCCAATTTCAA	438				
839	AGGCAAGTTCACCTTACTACGAATGATGCCATTAAAAACAACCTGTGCTGTGATGAATTTGGACA	898				
439	AGGCAAGTTCACCTTACTACGAATGATGCCATTAAAAACAACCTGTGCTGTGATGAATTTGGACA	498				
899	TTTCA 903					
499	TTTCA 503					

```

RESULT 14
US-09-949-016-150041
; Sequence 150041, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 150041
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-150041

```

Query Match 5.8%; Score 121; DB 4; Length 601;



Best Local Similarity 100.0%; Pred. No. 1.2e-51;  
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1556 AGATATCCATCTCTCTCGAACAACAATCTTTCACATTACAGATGACCCCTCAATCC 1615  
DB 124 AGATATCCATCTCTCTCGAACAACAATCTTTCACATTACAGATGACCCCTCAATCC 183

QY 1616 CCATCATTAATGTTGGGTCCAGAAACCGGCATAGCCCGTTATTGGGTCTTCAACATA 1675  
DB 184 CCATCATTAATGTTGGGTCCAGAAACCGGCATAGCCCGTTATTGGGTCTTCAACATA 243

QY 1676 G 1676  
DB 244 G 244

## RESULT 15

US-09-949-016-150042  
Sequence 150042, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CU001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 150042  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-150042

Query Match 5.8%; Score 121; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 1.2e-51;  
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1556 AGATATCCATCTCTCTCGAACAACAATCTTTCACATTACAGATGACCCCTCAATCC 1615  
DB 95 AGATATCCATCTCTCTCGAACAACAATCTTTCACATTACAGATGACCCCTCAATCC 154

QY 1616 CCATCATTAATGTTGGGTCCAGAAACCGGCATAGCCCGTTATTGGGTCTTCAACATA 1675  
DB 155 CCATCATTAATGTTGGGTCCAGAAACCGGCATAGCCCGTTATTGGGTCTTCAACATA 214

QY 1676 G 1676  
DB 215 G 215

Search completed: August 27, 2005, 16:18:16  
Job time : 237.757 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2005, 00:17:56 ; Search time 900.401 Seconds  
(without alignments)  
15238.608 Million cell updates/sec

Title: US-09-371-347a-41

Perfect score: 2097  
Sequence: 1 atgaggaggttctgttact.....ttcagagatattggtcacaata 2097

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 7331713 seqs, 3271544945 residues

Word size : 0

Total number of hits satisfying chosen parameters: 14663426

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 19: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 20: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 21: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 22: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 23: /cgn2\_6/ptodata/2/pubpna/US11\_PUBCOMB.seq:\*
- 24: /cgn2\_6/ptodata/2/pubpna/US11\_PUBCOMB.seq:\*
- 25: /cgn2\_6/ptodata/2/pubpna/US11\_PUBCOMB.seq:\*
- 26: /cgn2\_6/ptodata/2/pubpna/US11\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2097	100.0	2097	10	US-09-371-347-41
2	2046	97.6	3259	10	US-09-371-347-24
3	1995	95.1	2097	10	US-09-371-347-43
4	1803	86.0	2094	10	US-09-371-347-45
5	1742	83.1	2093	10	US-09-371-347-47
6	1062	50.6	3256	21	US-10-741-600-692

8	1062	50.6	3274	21	US-10-741-600-693	Sequence 693, App
9	330	15.7	591	16	US-10-029-386-6369	Sequence 6369, App
10	328	15.6	379	16	US-10-029-386-60100	Sequence 20100, App
11	279	13.3	591	16	US-10-029-386-1735	Sequence 1735, App
12	277	13.2	379	21	US-10-029-386-15435	Sequence 15435, App
13	266	12.7	43985	16	US-10-741-600-17757	Sequence 17757, App
14	188	9.0	525	16	US-10-029-386-633	Sequence 633, App
15	175	8.3	175	16	US-10-029-386-14338	Sequence 14338, App
16	158	7.5	2475	22	US-09-909-5678-38	Sequence 38, App
17	158	7.5	2475	22	US-10-765-700-88	Sequence 88, App
18	158	7.5	21852	21	US-10-741-600-17986	Sequence 17986, App
19	150	7.2	201	21	US-10-741-600-15583	Sequence 15583, App
20	150	7.2	201	21	US-10-741-600-15584	Sequence 15584, App
21	150	7.2	201	21	US-10-741-600-15589	Sequence 15589, App
22	150	7.2	201	21	US-10-741-600-15590	Sequence 15590, App
23	150	7.2	201	21	US-10-741-600-15592	Sequence 15592, App
24	150	7.2	201	21	US-10-741-600-15593	Sequence 15593, App
25	150	7.2	201	21	US-10-741-600-15594	Sequence 15594, App
26	150	7.2	201	21	US-10-741-600-15598	Sequence 15598, App
27	150	7.2	201	21	US-10-741-600-15599	Sequence 15599, App
28	150	7.2	201	21	US-10-741-600-15600	Sequence 15600, App
29	150	7.2	201	21	US-10-741-600-15602	Sequence 15602, App
30	150	7.2	201	21	US-10-741-600-15606	Sequence 15606, App
31	150	7.2	201	21	US-10-741-600-15609	Sequence 15609, App
32	150	7.2	201	21	US-10-741-600-15610	Sequence 15610, App
33	150	7.2	201	21	US-10-741-600-15612	Sequence 15612, App
34	150	7.2	201	21	US-10-741-600-15613	Sequence 15613, App
35	150	7.2	201	21	US-10-741-600-15614	Sequence 15614, App
36	150	7.2	201	21	US-10-741-600-15620	Sequence 15620, App
37	150	7.2	201	21	US-10-741-600-15621	Sequence 15621, App
38	150	7.2	201	21	US-10-741-600-15623	Sequence 15623, App
39	150	7.2	201	21	US-10-741-600-15624	Sequence 15624, App
40	150	7.2	201	21	US-10-741-600-15625	Sequence 15625, App
41	150	7.2	201	21	US-10-741-600-15629	Sequence 15629, App
42	150	7.2	201	21	US-10-741-600-15630	Sequence 15630, App
43	150	7.2	201	21	US-10-741-600-15631	Sequence 15631, App
44	150	7.2	201	21	US-10-741-600-15633	Sequence 15633, App
45	150	7.2	201	21	US-10-741-600-15637	Sequence 15637, App

#### ALIGNMENTS

RESULT 1  
US-09-371-347-41  
; Sequence 41, Application US/09371347  
; Publication No. US20030082676A1  
; GENERAL INFORMATION:  
; APPLICANT: Roy A. Gravel et al.  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE REDUCTASE;  
; TITLE OF INVENTION: CLONING, AND METHODS FOR EVALUATING RISK OF NEURAL TUBE  
; TITLE OF INVENTION: DEFECTS, CARDIOVASCULAR DISEASE, AND CANCER  
; FILE REFERENCE: 50004/003003  
; CURRENT APPLICATION NUMBER: US/09/371,347  
; CURRENT FILING DATE: 1999-08-10  
; PRIOR APPLICATION NUMBER: 60/071,622  
; PRIOR FILING DATE: 1998-01-16  
; PRIOR APPLICATION NUMBER: 09/232,028  
; PRIOR FILING DATE: 1999-01-15  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 41  
; LENGTH: 2097  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-371-347-41

Query Match 100.0%; Score 2097; DB 10; Length 2097;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 ATGAGGAGGTTCTGTACTATATGTCACAGCAGGAGCAGGCAAGGCATCGCAGAA 60  
|||||

Dh 1 ATGAGAGGTTTCTGTACTATATGCTACACAGAGGAGACAGCAAAAGCCATCGCAGAA 60  
Qy 61 GAATATGTAGCAAGCTGTGTATCATGATTTTCTGCAGATCTTCACTGATATAGTAA 120  
Dh 61 GAATATGTAGCAAGCTGTGTATCATGATTTTCTGCAGATCTTCACTGATATAGTAA 120  
Qy 121 TCCGATTAAGTATGACTTAAACCGAAACAGCTCTCTTGTGTGTGTTTCTACACG 180  
Dh 121 TCCGATTAAGTATGACTTAAACCGAAACAGCTCTCTTGTGTGTGTTTCTACACG 180  
Qy 181 GGACCGGAGAGCCACCGCAGACAGCCGCAAGTTTGTAAAGAAATACAAACCAACA 240  
Dh 181 GGACCGGAGAGCCACCGCAGACAGCCGCAAGTTTGTAAAGAAATACAAACCAACA 240  
Qy 241 CTGCGGTTGATTTCTTGTCTCACTGCGGTATGGTTACTGGTCTCGGTGATTCAGAA 300  
Dh 241 CTGCGGTTGATTTCTTGTCTCACTGCGGTATGGTTACTGGTCTCGGTGATTCAGAA 300  
Qy 301 TACACCTACTTTTGCATTTGGGGGAGATTAATTAAGACTTCAAGCTTGGAGCC 360  
Dh 301 TACACCTACTTTTGCATTTGGGGGAGATTAATTAAGACTTCAAGCTTGGAGCC 360  
Qy 361 CGGCAATTTCTATGACACTGACATGACAGATGACTGTAGGTTTGAACCTTGTGAG 420  
Dh 361 CGGCAATTTCTATGACACTGACATGACAGATGACTGTAGGTTTGAACCTTGTGAG 420  
Qy 421 CCGTGAATTTGCTGACCTTGCGCAGCCCTCAGAAAGATTTTATGTCAGACAGAGCAA 480  
Dh 421 CCGTGAATTTGCTGACCTTGCGCAGCCCTCAGAAAGATTTTATGTCAGACAGAGCAA 480  
Qy 481 GAGGAGATTAAGTGGCGCACTCCGCGTGGCATCACTGCTGCTTGAAGACAGACTTGTG 540  
Dh 481 GAGGAGATTAAGTGGCGCACTCCGCGTGGCATCACTGCTGCTTGAAGACAGACTTGTG 540  
Qy 541 AAGTCAGAGCTGCTACACATTTGTAAGCAAAATGCAAGTTCGAGATTCAGATTCAGGA 600  
Dh 541 AAGTCAGAGCTGCTACACATTTGTAAGCAAAATGCAAGTTCGAGATTCAGATTCAGGA 600  
Qy 601 AGAAAGATTTCTGAGGTTTGAACCAAAATGCAAGTTCAGATTCAGATTCAGGA 660  
Dh 601 AGAAAGATTTCTGAGGTTTGAACCAAAATGCAAGTTCAGATTCAGATTCAGGA 660  
Qy 661 ATTGAAGATTTGAGTCTCACTTACCCGTTCCGTTACCCCACTCTCAAGCTCTCTG 720  
Dh 661 ATTGAAGATTTGAGTCTCACTTACCCGTTCCGTTACCCCACTCTCTCAAGCTCTCTG 720  
Qy 721 AATATTTCTGCTTAAACCCCAAGATTTTACAGGTAATCTGACAGAGTCTCTTGGCCAG 780  
Dh 721 AATATTTCTGCTTAAACCCCAAGATTTTACAGGTAATCTGACAGAGTCTCTTGGCCAG 780  
Qy 781 GAGGAAAGCCAAAGTATCTGTGACTTCAGCAGATCCAGTTTTCAGTCCAAATTCAGAG 840  
Dh 781 GAGGAAAGCCAAAGTATCTGTGACTTCAGCAGATCCAGTTTTCAGTCCAAATTCAGAG 840  
Qy 841 GCAATTTAACTTACTAAGATGATGCAATTAACCACTCTGCTGTGATTAAGTGAATTC 900  
Dh 841 GCAATTTAACTTACTAAGATGATGCAATTAACCACTCTGCTGTGATTAAGTGAATTC 900  
Qy 901 TCAATATCAGACTTTTCTATCAGCTGAGATGCTTCAAGCTGATCTGCGCTTCAAGT 960  
Dh 901 TCAATATCAGACTTTTCTATCAGCTGAGATGCTTCAAGCTGATCTGCGCTTCAAGT 960  
Qy 961 GATTTCTGAGGTACAAAGCTTCTCAAGACTGACGTTGAGATTAAGAGAGAGACTGCG 1020  
Dh 961 GATTTCTGAGGTACAAAGCTTCTCAAGACTGACGTTGAGATTAAGAGAGAGACTGCG 1020  
Qy 1021 GTCTTTTGAATAATTAAGGAGACACAAAGAAAGAGAGTACCTTACCCACAGATTA 1080  
Dh 1021 GTCTTTTGAATAATTAAGGAGAGACACAAAGAAAGAGAGTACCTTACCCACAGATTA 1080  
Qy 1081 CTTGCGGAGATGTTCTCTCAGATTTTAACTGCTGCTGATTAAGTCCGAGCAATTCCT 1140  
Dh 1081 CTTGCGGAGATGTTCTCTCAGATTTTAACTGCTGCTGATTAAGTCCGAGCAATTCCT 1140

Qy 1141 AAAAAGCAATTTTGGAGCCCTTGTGACTATACAGTGAAGTCTGAAAAAGCCAGG 1200  
Dh 1141 AAAAAGCAATTTTGGAGCCCTTGTGACTATACAGTGAAGTCTGAAAAAGCCAGG 1200  
Qy 1201 CTACAGAGCTGTGCAAGTAAACAAAGGAGAGCCGATTAATAGCCGTTTGTACAGATGCG 1260  
Dh 1201 CTACAGAGCTGTGCAAGTAAACAAAGGAGAGCCGATTAATAGCCGTTTGTACAGATGCG 1260  
Qy 1261 TGTGCTGCTGTGTGATCTCTCTCTGCTTTCCTTCTTCCAGGACCACTCAAGTCTC 1320  
Dh 1261 TGTGCTGCTGTGTGATCTCTCTCTGCTTTCCTTCTTCCAGGACCACTCAAGTCTC 1320  
Qy 1321 CTGCTGAAATCTTCTTAACTTCAACCCAGACATTTGTGTGCAAGCTCAAGTTA 1380  
Dh 1321 CTGCTGAAATCTTCTTAACTTCAACCCAGACATTTGTGTGCAAGCTCAAGTTA 1380  
Qy 1381 TTTACCCAGAAAGCTCCATTTTGTCTTCAACATTTGTGGAATTTCTGCTACTGCGACA 1440  
Dh 1381 TTTACCCAGAAAGCTCCATTTTGTCTTCAACATTTGTGGAATTTCTGCTACTGCGACA 1440  
Qy 1441 ACAGAGTTCTGCGAAGGAGATGATGACAGCTGCGCTGTGTGTGCTTCAAGTT 1500  
Dh 1441 ACAGAGTTCTGCGAAGGAGATGATGACAGCTGCGCTGTGTGTGCTTCAAGTT 1500  
Qy 1501 CTTGAGCCAAACATACATGATCCATGAAAGACAGCGGAAAGCCCTGCTCTTAAGATA 1560  
Dh 1501 CTTGAGCCAAACATACATGATCCATGAAAGACAGCGGAAAGCCCTGCTCTTAAGATA 1560  
Qy 1561 TCCATCTCTCTCGAACAACAATTTCTTCCATTAACAGATGACCCCTCAATCCCATTC 1620  
Dh 1561 TCCATCTCTCTCGAACAACAATTTCTTCCATTAACAGATGACCCCTCAATCCCATTC 1620  
Qy 1621 ATATAGTGGGTTCAGAAACCGGATAGCCCTTATTTGGGTTCTTACAACTAGAGAG 1680  
Dh 1621 ATATAGTGGGTTCAGAAACCGGATAGCCCTTATTTGGGTTCTTACAACTAGAGAG 1680  
Qy 1681 AAATCTCAAGAACCAACCCAGATGAAATTTTGAAGCAATGTGTTTGTGCTGC 1740  
Dh 1681 AAATCTCAAGAACCAACCCAGATGAAATTTTGAAGCAATGTGTTTGTGCTGC 1740  
Qy 1741 AGGCATTAAGATGAGATTAATCTATTCAGAAAGACTCAGATTTCTTAAAGCATGGG 1800  
Dh 1741 AGGCATTAAGATGAGATTAATCTATTCAGAAAGACTCAGATTTCTTAAAGCATGGG 1800  
Qy 1801 ATCTTAATCATTAAGGTTTCTTCTCAAGAGATGCTCTGTGGGAGAGAGAGGCC 1860  
Dh 1801 ATCTTAATCATTAAGGTTTCTTCTCAAGAGATGCTCTGTGGGAGAGAGAGGCC 1860  
Qy 1861 CCAAGCAAGTATGTACAGACAAATCCAGCTTCATGCGCAGAGTGGCGAATCTTC 1920  
Dh 1861 CCAAGCAAGTATGTACAGACAAATCCAGCTTCATGCGCAGAGTGGCGAATCTTC 1920  
Qy 1921 CTCAGAGAAAGGCGCATTTTATGTGTGAGATGCAAGAAATATGGCCAGAGATGTA 1980  
Dh 1921 CTCAGAGAAAGGCGCATTTTATGTGTGAGATGCAAGAAATATGGCCAGAGATGTA 1980  
Qy 1981 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGTGTGAAGTTAAAAAATCAGAGCATG 2040  
Dh 1981 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGTGTGAAGTTAAAAAATCAGAGCATG 2040  
Qy 2041 AAAACCTGGCCACTTTAAAAAGAAAGAAAGCGTACCTTCAGATATTTGTCATTA 2097  
Dh 2041 AAAACCTGGCCACTTTAAAAAGAAAGAAAGCGTACCTTCAGATATTTGTCATTA 2097

RESULT 2  
US-09-371-347-1  
; Sequence 1, Application US/09371347  
; Publication No. US20030082676A1  
; GENERAL INFORMATION:  
; APPLICANT: Roy A. Gravel et al.  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE REDUCTASE;

TITLE OF INVENTION: CLONING, AND METHODS FOR EVALUATING RISK OF NEURAL TUBE  
DEFECTS CARDIOVASCULAR DISEASE, AND CANCER  
FILE REFERENCE: 50004/003003  
CURRENT APPLICATION NUMBER: US/09/371,347  
CURRENT FILING DATE: 1999-08-10  
PRIOR APPLICATION NUMBER: 60/071,622  
PRIOR FILING DATE: 1998-01-16  
PRIOR APPLICATION NUMBER: 09/232,028  
PRIOR FILING DATE: 1999-01-15  
NUMBER OF SEQ ID NOS: 51  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO: 1  
LENGTH: 2097  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-371-347-1

Query Match 97.6%; Score 2046; DB 10; Length 2097;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2096; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGAGGTTCTGTACTATATGCTACACAGCAGGAGGAGGCAAGGCCATCGCAGAA 60  
DB 1 ATGAGAGGTTCTGTACTATATGCTACACAGCAGGAGGAGGCAAGGCCATCGCAGAA 60  
QY 61 GAAATATGTGAGCAGAGCTGTGTACATGATTTTCTGAGATCTTCACTGATTAAGAA 120  
DB 61 GAAATATGTGAGCAGAGCTGTGTACATGATTTTCTGAGATCTTCACTGATTAAGAA 120  
QY 121 TCCGATATGATGACCTTAATAACCGAAAGAGCTCTCTGTGTGTGTGTGTGTCTACAG 180  
DB 121 TCCGATATGATGACCTTAATAACCGAAAGAGCTCTCTGTGTGTGTGTGTGTCTACAG 180  
QY 181 GGCACCGGAGACCCACCGCAGACAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACA 240  
DB 181 GGCACCGGAGACCCACCGCAGACAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACA 240  
QY 241 CTGCGGTTGATTTCTTGTCTCACTGCGGTATGAGGTCTCGGTATTCAGAA 300  
DB 241 CTGCGGTTGATTTCTTGTCTCACTGCGGTATGAGGTCTCGGTATTCAGAA 300  
QY 301 TACACCTACTTTTGAATGGGGGAGATTAATTGATTAACGACTTCAAGAGCTTGAGCC 360  
DB 301 TACACCTACTTTTGAATGGGGGAGATTAATTGATTAACGACTTCAAGAGCTTGAGCC 360  
QY 361 CGGCAATTTCTATGACACTGGAATGCAATGATGATGATGATGATGATGATGATGATG 420  
DB 361 CGGCAATTTCTATGACACTGGAATGCAATGATGATGATGATGATGATGATGATGATG 420  
QY 421 CCGTGATGTTGCTGAGCTCTGCGCAGCCCTCAGAAAGATTTTATGATGATGATGATG 480  
DB 421 CCGTGATGTTGCTGAGCTCTGCGCAGCCCTCAGAAAGATTTTATGATGATGATGATG 480  
QY 481 GAGGAGATTAAGTGGGCACTCCGCGTGGGATCACTGATCTTGAAGGACAGAGCTTGTG 540  
DB 481 GAGGAGATTAAGTGGGCACTCCGCGTGGGATCACTGATCTTGAAGGACAGAGCTTGTG 540  
QY 541 AAGTCAGAGCTGCTACATTAAGTGAATCTCAAGTGAAGCTTCTGAGATTCAGATTCAG 600  
DB 541 AAGTCAGAGCTGCTACATTAAGTGAATCTCAAGTGAAGCTTCTGAGATTCAGATTCAG 600  
QY 601 AGAAGAGATTTCTGAGGTTTGAAGCAAAATGCAAGTGAAGCAAGCAATCCAAATGTTGTA 660  
DB 601 AGAAGAGATTTCTGAGGTTTGAAGCAAAATGCAAGTGAAGCAAGCAATCCAAATGTTGTA 660  
QY 661 ATGAGAGCTTTGAGTCTCACTTACCCGTTGGTACCCCACTCTCAAGAGCTCTG 720  
DB 661 ATGAGAGCTTTGAGTCTCACTTACCCGTTGGTACCCCACTCTCAAGAGCTCTG 720  
QY 721 AATATTCCTGTTTACCCCAAGATATTTTACAGATCACTGCAAGAGTCTTGGCCAG 780  
DB 721 AATATTCCTGTTTACCCCAAGATATTTTACAGATCACTGCAAGAGTCTTGGCCAG 780

QY 781 GAGGAAAGCAAGTATCTGTGACTTCAAGAGATTCAGATTTTCAAGTGCATTTCAAG 840  
DB 781 GAGGAAAGCAAGTATCTGTGACTTCAAGAGATTCAGATTTTCAAGTGCATTTCAAG 840  
QY 841 GCAATTCATCTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 900  
DB 841 GCAATTCATCTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 900  
QY 901 TCAATTCAGATTTTCTTCAAGGCTGAGATGCTTCAAGGATGATGATGATGATGATG 960  
DB 901 TCAATTCAGATTTTCTTCAAGGCTGAGATGCTTCAAGGATGATGATGATGATGATG 960  
QY 961 GATTCAGAGTACCAAGCTTACCAAGAGCTGAGATGATGATGATGATGATGATGATG 1020  
DB 961 GATTCAGAGTACCAAGCTTACCAAGAGCTGAGATGATGATGATGATGATGATGATG 1020  
QY 1021 GTCTTTTGAATAAAGGAGACACAAAGAAAGAGGATGATGATGATGATGATGATG 1080  
DB 1021 GTCTTTTGAATAAAGGAGACACAAAGAAAGAGGATGATGATGATGATGATGATG 1080  
QY 1081 CCGGAGGATGCTTCTCAAGTCTTCTTCAAGTCTTCTTCAAGTCTTCTTCAAGTCT 1140  
DB 1081 CCGGAGGATGCTTCTCAAGTCTTCTTCAAGTCTTCTTCAAGTCTTCTTCAAGTCT 1140  
QY 1141 AAAAAAGCATTTTTCAGAGCCCTTGTGAGCTATACAGTGAAGTGTGAAAAAGCCAG 1200  
DB 1141 AAAAAAGCATTTTTCAGAGCCCTTGTGAGCTATACAGTGAAGTGTGAAAAAGCCAG 1200  
QY 1201 CTACAGAGCTGTGAGTAAACAGAGGAGCCGATTAAGCCGCTTGTGAGAGATG 1260  
DB 1201 CTACAGAGCTGTGAGTAAACAGAGGAGCCGATTAAGCCGCTTGTGAGAGATG 1260  
QY 1261 TGTGCTGCTTGTGATCTCTCTCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 1320  
DB 1261 TGTGCTGCTTGTGATCTCTCTCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 1320  
QY 1321 CTGCTGGAACATCTTCTTAACTTCAACCCAGACATATTCGTGCAAGCTCAAGTTTA 1380  
DB 1321 CTGCTGGAACATCTTCTTAACTTCAACCCAGACATATTCGTGCAAGCTCAAGTTTA 1380  
QY 1381 TTTCAACCCAGAAAGCTCAATTTTGTCTTCAATTTGTGAAATTTCTGTCTACCTG 1440  
DB 1381 TTTCAACCCAGAAAGCTCAATTTTGTCTTCAATTTGTGAAATTTCTGTCTACCTG 1440  
QY 1441 ACAAGGTTCTGCGAAGGAGATGATGATGATGATGATGATGATGATGATGATGAT 1500  
DB 1441 ACAAGGTTCTGCGAAGGAGATGATGATGATGATGATGATGATGATGATGATGAT 1500  
QY 1501 CTTCAAGCCAAATACATGATCCCATGAAAGCAGCGGAGAAAGCCCTGAGCTCTAAGATA 1560  
DB 1501 CTTCAAGCCAAATACATGATCCCATGAAAGCAGCGGAGAAAGCCCTGAGCTCTAAGATA 1560  
QY 1561 TCCATCTCTCTGAGCAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCCTAT 1620  
DB 1561 TCCATCTCTCTGAGCAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCCTAT 1620  
QY 1621 ATATATGTTGTTCTGAGAAACCGGATGATGATGATGATGATGATGATGATGATG 1680  
DB 1621 ATATATGTTGTTCTGAGAAACCGGATGATGATGATGATGATGATGATGATGATG 1680  
QY 1681 AATCTCAAGAAACACACAGATGAAATTTTGAAGCAATGATGATGATGATGATGATG 1740  
DB 1681 AATCTCAAGAAACACACAGATGAAATTTTGAAGCAATGATGATGATGATGATGATG 1740  
QY 1741 AGGCAATTAAGATAGGATTAATCTTCAAGAAAGAGCTCAGATTTCTTAAAGATGAG 1800  
DB 1741 AGGCAATTAAGATAGGATTAATCTTCAAGAAAGAGCTCAGATTTCTTAAAGATGAG 1800  
QY 1801 ATCTTAATCTTAAGATTTCTTCTCAAGATGCTCTGTTGGGAGAGAGAGAGCC 1860  
DB 1801 ATCTTAATCTTAAGATTTCTTCTCAAGATGCTCTGTTGGGAGAGAGAGAGCC 1860  
QY 1861 CCAGCAAGATATGATCAAGCAACATCAGCTTCAATGTCAGAGATGTCAGAAATCTCT 1920

```
Db      1861 CCGAGCAAGATGATGACAAAGCAACATCAGCTTCATGCGACAGAGGGGAGGAATCTC 1920
Qy      1921 CTCAGAGAGACGGCCATATTTATGTGTGAGATCAAGATATATGCGCAAGATGTA 1980
Db      1921 CTCAGAGAGACGGCCATATTTATGTGTGAGATCAAGATATATGCGCAAGATGTA 1980
Qy      1981 CATGATCCCTTGTGCAAAATATATTAAGCAAGAGGTTGAGTTGAAAACTAGAAGCATG 2040
Db      1981 CATGATCCCTTGTGCAAAATATATTAAGCAAGAGGTTGAGTTGAAAACTAGAAGCATG 2040
Qy      2041 AAAACCTGGCCACTTTAAAAAGAAAGAAACGCTACCTTCAGATATTTGTCATTA 2097
Db      2041 AAAACCTGGCCACTTTAAAAAGAAAGAAACGCTACCTTCAGATATTTGTCATTA 2097

RESULT 3
US-09-371-347-24
; Sequence 24, Application US/09371347
; Publication No. US20030082676A1
; GENERAL INFORMATION:
; APPLICANT: Roy A. Gravel et al.
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE REDUCTASE:
; TITLE OF INVENTION: CLONING, AND METHODS FOR EVALUATING RISK OF NEURAL TUBE
; TITLE OF INVENTION: DEFECTS, CARDIOVASCULAR DISEASE, AND CANCER
; FILE REFERENCE: 50004/003003
; CURRENT APPLICATION NUMBER: US/09/371,347
; CURRENT FILING DATE: 1999-08-10
; PRIOR APPLICATION NUMBER: 60/071,622
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 09/232,028
; PRIOR FILING DATE: 1999-01-15
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: FaSTSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 3259
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-371-347-24

Query Match      97.6%; Score 2046; DB 10; Length 3259;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2096; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 ATGAGAGGTTTCTGTACTATATGCTACACAGCAGGAGACAGCAAGCCATGCGAGAA 60
Db      80 ATGAGAGGTTTCTGTACTATATGCTACACAGCAGGAGACAGCAAGCCATGCGAGAA 139
Qy      61 GAAATATGTAGAGAGCTGTGTGATCATGATTTTCTGACATCTTCACTGATTAAGTAA 120
Db      140 GAAATGTGTAGAGAGCTGTGTGATCATGATTTTCTGACATCTTCACTGATTAAGTAA 199
Qy      121 TCCGATATGATGACTTAAAAACGAAACAGCTCCTGTGTGTGTGTGTGTTTCTACCAAG 180
Db      200 TCCGATATGATGACTTAAAAACGAAACAGCTCCTGTGTGTGTGTGTGTTTCTACCAAG 259
Qy      181 GGCACCGGAGACCCACCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACA 240
Db      260 GGCACCGGAGACCCACCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACA 319
Qy      241 CTGCGCGGTGATTTCTTTGTCTACACCTGCGGTGATAGGTTTCTGCGTCTCGGATTCAGAA 300
Db      320 CTGCGCGGTGATTTCTTTGTCTACACCTGCGGTGATAGGTTTCTGCGTCTCGGATTCAGAA 379
Qy      301 TACACCTACTTTTTCGAATGGGGGAGATTAATTGATTAACGACTTCAAGAGCTTGGAGCC 360
Db      380 TACACCTACTTTTTCGAATGGGGGAGATTAATTGATTAACGACTTCAAGAGCTTGGAGCC 439
Qy      361 CGGCAATTTCTATGACATGACATGACAGATGACTGTAGGTTTGAACCTTGTGTTGAG 420
Db      440 CGGCAATTTCTATGACATGACATGACAGATGACTGTAGGTTTGAACCTTGTGTTGAG 499
Qy      421 CCGTGATGTGCTGGAATCTGGCCAGCCCTCAGAAAGCATTTTAAAGTCAAGAGAGAGCAA 480
```

```
Db      500 CCGTGATGTGCTGGAATCTGGCCAGCCCTCAGAAAGCATTTTAAAGTCAAGAGAGAGCAA 559
Qy      481 GAGGAGATAGTGGCCGCACTCCGGTGGATCATCTGCATCTTTGAGGAGACAGACTTGTG 540
Db      560 GAGGAGATAGTGGCCGCACTCCGGTGGATCATCTGCATCTTTGAGGAGACAGACTTGTG 619
Qy      541 AAGTCAGACTGTGACACATTTGATTCATAGTCGAGCTTCTGAGATTCGATGATTCAGGA 600
Db      620 AAGTCAGACTGTGACACATTTGATTCATAGTCGAGCTTCTGAGATTCGATGATTCAGGA 679
Qy      601 AGAAGAGATTCGAGGTTTGAAGCAAAATGCAAGTAAAGCAAGCAATCCAAATGTTGTA 660
Db      680 AGAAGAGATTCGAGGTTTGAAGCAAAATGCAAGTAAAGCAAGCAATCCAAATGTTGTA 739
Qy      661 ATTGAAGCTTTGAGTCTCATCTTACCCGTTGGATACCCCACTTCACAGGCTCTG 720
Db      740 ATTGAAGCTTTGAGTCTCATCTTACCCGTTGGATACCCCACTTCACAGGCTCTG 799
Qy      721 AATATTCCTGTTTACCCCAAGATTTTACAGGTACATCTGACAGAGTCTTGGCCAG 780
Db      800 AATATTCCTGTTTACCCCAAGATTTTACAGGTACATCTGACAGAGTCTTGGCCAG 859
Qy      781 GAGAAAGCCAGATCTGTGACTTCGACAGATCCAGTTTTCAGTGGCCAAATTTGAAAG 840
Db      860 GAGAAAGCCAGATCTGTGACTTCGACAGATCCAGTTTTCAGTGGCCAAATTTGAAAG 919
Qy      841 GCAGTTCAACTTACATGAGATGATGCAATAAACCACTCTGCTGTAGATTTGACATT 900
Db      920 GCAGTTCAACTTACATGAGATGATGCAATAAACCACTCTGCTGTAGATTTGACATT 979
Qy      901 TCAGATACAGACTTTTCTATCAGCTCGAGATGCTTCAAGCTGATCTGCCCTAACAGT 960
Db      980 TCAGATACAGACTTTTCTATCAGCTCGAGATGCTTCAAGCTGATCTGCCCTAACAGT 1039
Qy      961 GATTCGAGGTACAAAGCCATCTCCAAAGCTGCACTTGAAGATTAAGAGAGCACTGC 1020
Db      1040 GATTCGAGGTACAAAGCCATCTCCAAAGCTGCACTTGAAGATTAAGAGAGCACTGC 1099
Qy      1021 GTCCCTTTGAAATTAAGGACGACACAAAGAAAGAGACTTACCTTACCCGACATATA 1080
Db      1100 GTCCCTTTGAAATTAAGGACGACACAAAGAAAGAGACTTACCTTACCCGACATATA 1159
Qy      1081 CTGCGGAGATGTTCTCTCCAGTTCATTTTCTGCTGTCTTGAATTCGAGCAATTTCT 1140
Db      1160 CTGCGGAGATGTTCTCTCCAGTTCATTTTCTGCTGTCTTGAATTCGAGCAATTTCT 1219
Qy      1141 AAAAAGGCAATTTTGGAGGCCCTGTGGACTATACAGAGTACAGTGTGAAAAGCGCAGG 1200
Db      1220 AAAAAGGCAATTTTGGAGGCCCTGTGGACTATACAGAGTACAGTGTGAAAAGCGCAGG 1279
Qy      1201 CTACAGAGCTGTGAGTAAACAAAGGGGACGCCGATTAAGCCGTTTGTACGAGATGCC 1260
Db      1280 CTACAGAGCTGTGAGTAAACAAAGGGGACGCCGATTAAGCCGTTTGTACGAGATGCC 1339
Qy      1261 TGTGCTGCTGTTGAGATCTCTCTGCTGCTTTCCCTTCTTGGCAGCACAATCACTCTC 1320
Db      1340 TGTGCTGCTGTTGAGATCTCTCTGCTGCTTTCCCTTCTTGGCAGCACAATCACTCTC 1399
Qy      1321 CTGCTGGAACATCTTCTTAACTTCAACCGACGACATATTGNTGNTGAGCTCAAGTTTA 1380
Db      1400 CTGCTGGAACATCTTCTTAACTTCAACCGACGACATATTGNTGNTGAGCTCAAGTTTA 1459
Qy      1381 TTTTCAACCGAAGAGCTCAATTTTGTCTTCAACATTTGTGAGATTTGTCTACTGCCACA 1440
Db      1460 TTTTCAACCGAAGAGCTCAATTTTGTCTTCAACATTTGTGAGATTTGTCTACTGCCACA 1519
Qy      1441 ACAGAGTTTCTGCGAAGGAGATATGATACAGGCTGCTGCTGCTTGTGTTGCTTCAATT 1500
Db      1520 ACAGAGTTTCTGCGAAGGAGATATGATACAGGCTGCTGCTGCTTGTGTTGCTTCAATT 1579
Qy      1501 CTTTCAAGCAAACTATACATGATCCCATGAAGACAGGGGGAAGCCCTGCTCTCAAGATA 1560
```

Db 1580 CTTGACCCAAACATCATGATCCCATGAAGACAGCGGAAAACCTGCTCTTAAGATA 1639  
Qy 1561 TCCATCTCTCTGAGAACAAATTCCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
Db 1640 TCCATCTCTCTGAGAACAAATTCCTTCACTTACAGATGACCCCTCAATCCCATC 1699  
Qy 1621 ATATATGTGGGTCAGAGAACCCGACATAGCCCGTTTATTTGGGTTCTTACACATAGAG 1680  
Db 1700 ATATATGTGGGTCAGAGAACCCGACATAGCCCGTTTATTTGGGTTCTTACACATAGAG 1759  
Qy 1681 AAACTCCAGAACACACACCCAGATGAAATTTTGGAGCAATGTGTGTTTGGCTGC 1740  
Db 1760 AAACTCCAGAACACACACCCAGATGAAATTTTGGAGCAATGTGTGTTTGGCTGC 1819  
Qy 1741 AGGCAATAGATATGGAATTAATCATTCAGAAAAGACTCAGACATTTCTTAAGATGG 1800  
Db 1820 AGGCAATAGATATGGAATTAATCATTCAGAAAAGACTCAGACATTTCTTAAGATGG 1879  
Qy 1801 ATCTTAATCATTAAGGTTTCTTCTCAGAGATGCTCTGTTGGGAGAGAGAGCC 1860  
Db 1880 ATCTTAATCATTAAGGTTTCTTCTCAGAGATGCTCTGTTGGGAGAGAGAGCC 1939  
Qy 1861 CCAGCAAGATATGATACAGACATTCAGCTTCATGCGCAGAGGTGGCAGATCTTC 1920  
Db 1940 CCAGCAAGATATGATACAGACATTCAGCTTCATGCGCAGAGGTGGCAGATCTTC 1999  
Qy 1921 CTCACAGAAACCGCATATTTATGTGTGTGAGATGCAAAAGATATGCGCAAGATTA 1980  
Db 2000 CTCACAGAAACCGCATATTTATGTGTGTGAGATGCAAAAGATATGCGCAAGATTA 2059  
Qy 1981 CATGATGCTCTGTGCAAAATTAATTAAGCAAGGTTGAGTTGAAAACTAGAAGCATG 2040  
Db 2060 CATGATGCTCTGTGCAAAATTAATTAAGCAAGGTTGAGTTGAAAACTAGAAGCATG 2119  
Qy 2041 AAAACCTGGCCACTTTAAAAAGAAAAACGCTACCTTCAGGATATTTGGTCATTA 2097  
Db 2120 AAAACCTGGCCACTTTAAAAAGAAAAACGCTACCTTCAGGATATTTGGTCATTA 2176

RESULT 4  
US-09-371-347-43  
; Sequence 43, Application US/09371347  
; Publication No. US2003082676A1  
; GENERAL INFORMATION:  
; APPLICANT: Roy A. Gravel et al.  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE REDUCTASE:  
; TITLE OF INVENTION: CLONING, AND METHODS FOR EVALUATING RISK OF NEURAL TUBE  
; FILE REFERENCE: 50004/003003  
; CURRENT APPLICATION NUMBER: US/09/371,347  
; CURRENT FILING DATE: 1999-08-10  
; PRIOR APPLICATION NUMBER: 60/071,622  
; PRIOR FILING DATE: 1998-01-16  
; PRIOR APPLICATION NUMBER: 09/232,028  
; PRIOR FILING DATE: 1999-01-15  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: FaastSeq for Windows Version 4.0  
; SEQ ID NO 43  
; LENGTH: 2097  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-371-347-43

Query Match 95.1%; Score 1995; DB 10; Length 2097;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2095; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATAGAGAGGTTCTGTTACTATATGCTACACAGAGGACAGGCAAGGCATCGCAGAA 60  
Db 1 ATAGAGAGGTTCTGTTACTATATGCTACACAGAGGACAGGCAAGGCATCGCAGAA 60  
Qy 61 GAAATATGTAGAGAGCTGTGTGATCATGATTTTCTGACATCTTCACTATATAGTAA 120  
Db 61 GAAATATGTAGAGAGCTGTGTGATCATGATTTTCTGACATCTTCACTATATAGTAA 120

Db 61 GAAATATGTAGAGAGCTGTGTGATCATGATTTTCTGACATCTTCACTATATAGTAA 120  
Qy 121 TCCGATTAAGTATGACCTTAAAAACGAAACAGCTCTCTGTGTGTGTTTCTACACAG 180  
Db 121 TCCGATTAAGTATGACCTTAAAAACGAAACAGCTCTCTGTGTGTGTTTCTACACAG 180  
Qy 181 GGCACCGAGAGACCCAGCCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCA 240  
Db 181 GGCACCGAGAGACCCAGCCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCA 240  
Qy 241 CTGCGGATGATTTCTTGTCTCACTGCGGTATAGGTTACTGGGTCTCGGTATTCAGA 300  
Db 241 CTGCGGATGATTTCTTGTCTCACTGCGGTATAGGTTACTGGGTCTCGGTATTCAGA 300  
Qy 301 TACACCTACTTTTGCATATGAGGAGAAAGATTAATGATTAACGACTTCAAGAGCTTGA 360  
Db 301 TACACCTACTTTTGCATATGAGGAGAAAGATTAATGATTAACGACTTCAAGAGCTTGA 360  
Qy 361 CGGCAATTTCTATGACACTGACATGACAGATGAATGTATAGTTTGAACCTTGTTGAG 420  
Db 361 CGGCAATTTCTATGACACTGACATGACAGATGAATGTATAGTTTGAACCTTGTTGAG 420  
Qy 421 CCGTGATGCTGGAATCTGCGCAGCCCTCAGAAAGCATTTTATAGTCAAGCAGAGACA 480  
Db 421 CCGTGATGCTGGAATCTGCGCAGCCCTCAGAAAGCATTTTATAGTCAAGCAGAGACA 480  
Qy 481 GAGGAGATAGTGGGCACTCCCGGTGGATACCGTGCATCTTGAGAGACAGACCTGTG 540  
Db 481 GAGGAGATAGTGGGCACTCCCGGTGGATACCGTGCATCTTGAGAGACAGACCTGTG 540  
Qy 541 AAGTCAGAGCTCTACACATTTGATCTCAATGAGCTTCTGAGATTCATGATTCAGAGA 600  
Db 541 AAGTCAGAGCTCTACACATTTGATCTCAATGAGCTTCTGAGATTCATGATTCAGAGA 600  
Qy 601 AGAAGAGATTCGAGGTTTGAAGCAAAATGCAATGAAACAGCAACCAATTCATGTTGA 660  
Db 601 AGAAGAGATTCGAGGTTTGAAGCAAAATGCAATGAAACAGCAACCAATTCATGTTGA 660  
Qy 661 ATTGAAGATTTGAGTCCCTCACTTACCGGTGGATACCCCACTGTCAGAAAGCTCTG 720  
Db 661 ATTGAAGATTTGAGTCCCTCACTTACCGGTGGATACCCCACTGTCAGAAAGCTCTG 720  
Qy 721 AATATTCCTGTTTACCCCGCAAAATTTTACAGATCATCTGAGAGTCTCTGGCCAG 780  
Db 721 AATATTCCTGTTTACCCCGCAAAATTTTACAGATCATCTGAGAGTCTCTGGCCAG 780  
Qy 781 GAGGAAAGCCAGATCTGTGACTTCAGAGATTCAGATTTTCAAGTGCATTTCAAG 840  
Db 781 GAGGAAAGCCAGATCTGTGACTTCAGAGATTCAGATTTTCAAGTGCATTTCAAG 840  
Qy 841 GCAATTCATTAATGATGCAATGATGCAATTAACCACTCTGTGTGAATTTGACATT 900  
Db 841 GCAATTCATTAATGATGCAATGATGCAATTAACCACTCTGTGTGAATTTGACATT 900  
Qy 901 TCAATATGACATTTTCTATGAGCTGTGAGATGCTTCAAGCTGATCTGCTTACAG 960  
Db 901 TCAATATGACATTTTCTATGAGCTGTGAGATGCTTCAAGCTGATCTGCTTACAG 960  
Qy 961 GATTCGAGGATCAAGAGCTTCTCAAGATGACAGCTTGAAGATTAAGAGAGCCTGC 1020  
Db 961 GATTCGAGGATCAAGAGCTTCTCAAGATGACAGCTTGAAGATTAAGAGAGCCTGC 1020  
Qy 1021 GTTCCTTTGAAATTAAGGACACACAAAGAAAGAGGCTTACCTTACCCGACATTA 1080  
Db 1021 GTTCCTTTGAAATTAAGGACACACAAAGAAAGAGGCTTACCTTACCCGACATTA 1080  
Qy 1081 CCGCGGAGATGTTCTCTCAGATTCATTTTACCTGATCTTGAATCCGAGCAATTCCT 1140  
Db 1081 CCGCGGAGATGTTCTCTCAGATTCATTTTACCTGATCTTGAATCCGAGCAATTCCT 1140  
Qy 1141 AAAAGGCAATTTTGGAGCCCTTGTGACTATACAGTACAGTGTCTGAAAAGGCCAGG 1200  
Db 1141 AAAAGGCAATTTTGGAGCCCTTGTGACTATACAGTACAGTGTCTGAAAAGGCCAGG 1200

QY 1201 CTACAGAGCTGTGACGATTAACAAAGGGGACCGGATTAATAGCCGCTTTGTAACGATGCC 1260  
DB 1201 CTACAGAGCTGTGACGATTAACAAAGGGGACCGGATTAATAGCCGCTTTGTAACGATGCC 1260  
QY 1261 TGTGACCTGCTGTGATCTCTCTGCTTTCCCTTGGCCAGCAGCACTGATCTC 1320  
DB 1261 TGTGACCTGCTGTGATCTCTCTGCTTTCCCTTGGCCAGCAGCACTGATCTC 1320  
QY 1321 CTGCTGACATCTTCTTAACTTCAACCCAGACATATTCGTGTGACAGCTCAAGTTTA 1380  
DB 1321 CTGCTGACATCTTCTTAACTTCAACCCAGACATATTCGTGTGACAGCTCAAGTTTA 1380  
QY 1381 TTTGACCCAGAGAGCTTCATTTTGTCTTCAACTTGTGAATTTCTGTCTACGCGACA 1440  
DB 1381 TTTGACCCAGAGAGCTTCATTTTGTCTTCAACTTGTGAATTTCTGTCTACGCGACA 1440  
QY 1441 ACAGAGGTTCTGCGAGAGGGAGTATGTAACGGCTGCGCTTGTGTGCTTCAAGT 1500  
DB 1441 ACAGAGGTTCTGCGAGAGGGAGTATGTAACGGCTGCGCTTGTGTGCTTCAAGT 1500  
QY 1501 CTTCAGCCAAACATACATGATGCCATGAAAGACCGGAAAGCCCTGCTCTTAAGATA 1560  
DB 1501 CTTCAGCCAAACATACATGATGCCATGAAAGACCGGAAAGCCCTGCTCTTAAGATA 1560  
QY 1561 TCCATCTCTCTGCAACACAAATTTCTTCACTTACAGATGACCCCTGAATCCCATC 1620  
DB 1561 TCCATCTCTCTGCAACACAAATTTCTTCACTTACAGATGACCCCTGAATCCCATC 1620  
QY 1621 ATATATGAGGTTCAGAGAACCGGCAATAGCCCGTTTATGAGGTTCCTTCAACATAGAG 1680  
DB 1621 ATATATGAGGTTCAGAGAACCGGCAATAGCCCGTTTATGAGGTTCCTTCAACATAGAG 1680  
QY 1681 AAATCTCAAGAAACAACCCAGATGAAATTTTGGAGCAATGTGTGTTTTTGGCTGC 1740  
DB 1681 AAATCTCAAGAAACAACCCAGATGAAATTTTGGAGCAATGTGTGTTTTTGGCTGC 1740  
QY 1741 AGGCATTAAGATAGGATTTATCTATTCAGAAAGAGCTCAGACATTTCTTAAAGCATG 1800  
DB 1741 AGGCATTAAGATAGGATTTATCTATTCAGAAAGAGCTCAGACATTTCTTAAAGCATG 1800  
QY 1801 ATCTTAATCTATCTTAAAGGTTTTCTTCTCAAGAGATCTCTGTGTGGGAGAGAGAGCC 1860  
DB 1801 ATCTTAATCTATCTTAAAGGTTTTCTTCTCAAGAGATCTCTGTGTGGGAGAGAGAGCC 1860  
QY 1861 CCAAGCAAGTATGTAACAAGACATCAGCTTCATGCGCAGCAGGTGGCAGATCTC 1920  
DB 1861 CCAAGCAAGTATGTAACAAGACATCAGCTTCATGCGCAGCAGGTGGCAGATCTC 1920  
QY 1921 CTCCAGAGAAAGGCAATTTATGTGTGTGAGATGCAAGAAATATGCGCAAGATGTA 1980  
DB 1921 CTCCAGAGAAAGGCAATTTATGTGTGTGAGATGCAAGAAATATGCGCAAGATGTA 1980  
QY 1981 CATATATCCCTTGTGCAATTAATAAGAAAGGTTGAGTTGAAAACTTAAAGCATG 2040  
DB 1981 CATATATCCCTTGTGCAATTAATAAGAAAGGTTGAGTTGAAAACTTAAAGCATG 2040  
QY 2041 AAAACCTGGCCCTTTAAAGAGAAAAAGCTACCTTCAGATTTTGGTCTAA 2097  
DB 2041 AAAACCTGGCCCTTTAAAGAGAAAAAGCTACCTTCAGATTTTGGTCTAA 2097

RESULT 5  
US-09-371-347-45  
; Sequence 45, Application us/09371347  
; Publication No. US20030082676A1  
; GENERAL INFORMATION:  
; APPLICANT: Roy A. Gravel et al.  
; TITLE OF INVENTION: HUMAN METHYLONINE SYNTHASE REDUCTASE;  
; TITLE OF INVENTION: CLONING AND METHODS FOR EVALUATING RISK OF NEURAL TUBE  
; TITLE OF INVENTION: DEFECTS, CARDIOVASCULAR DISEASE, AND CANCER  
; FILE REFERENCE: 50004/003003  
; CURRENT APPLICATION NUMBER: US/09/371.347

; CURRENT FILING DATE: 1999-08-10  
; PRIOR APPLICATION NUMBER: 60/071,622  
; PRIOR FILING DATE: 1998-01-16  
; PRIOR APPLICATION NUMBER: 09/232,028  
; PRIOR FILING DATE: 1999-01-15  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 45  
; LENGTH: 2094  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-371-347-45

Query Match 86.0%; Score 1803; DB 10; Length 2094;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2093; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 1 ATGAGAGGTTCTGTACTATATGCTACACAGCAGGACAGGCAAGGCATCGCAGAA 60  
DB 1 ATGAGAGGTTCTGTACTATATGCTACACAGCAGGACAGGCAAGGCATCGCAGAA 60  
QY 61 GAAATATGTAGCAGACCTGTGTACATGATTTTCTGCAATCTTCACTGTATTAGTGA 120  
DB 61 GAAATATGTAGCAGACCTGTGTACATGATTTTCTGCAATCTTCACTGTATTAGTGA 120  
QY 121 TCCGATTAAGTATGACTTAAACCGAAACAGCTCTCTTGTGTGTGTTTCTTACAG 180  
DB 121 TCCGATTAAGTATGACTTAAACCGAAACAGCTCTCTTGTGTGTGTTTCTTACAG 180  
QY 181 GGACAGGAGAACCCACCGACACAGCCCGCAAGTTGTTAAGAAATPACAGAACAA 240  
DB 181 GGACAGGAGAACCCACCGACACAGCCCGCAAGTTGTTAAGAAATPACAGAACAA 240  
QY 241 CTGCGGTTGATTTCTTGTCTCACTGCGATAGGTTTCTGAGTCTCGGTATTCAGAA 300  
DB 241 CTGCGGTTGATTTCTTGTCTCACTGCGATAGGTTTCTGAGTCTCGGTATTCAGAA 300  
QY 301 TACACCTACTTTTGCATGAGGAGGAAATATGATTAACGACTTCAAGAGCTTGAACC 360  
DB 301 TACACCTACTTTTGCATGAGGAGGAAATATGATTAACGACTTCAAGAGCTTGAACC 360  
QY 361 CGGCAATTTCTATGACACTGACATGAGATGATGATGATGATGATGATGATGATGAT 420  
DB 361 CGGCAATTTCTATGACACTGACATGAGATGATGATGATGATGATGATGATGATGAT 420  
QY 421 CCGTGATTTGCTGACTGTGCGACGCTCAGAAACATTTTAAAGTCAAGCAGAGACA 480  
DB 421 CCGTGATTTGCTGACTGTGCGACGCTCAGAAACATTTTAAAGTCAAGCAGAGACA 480  
QY 481 GAGAGATTAAGTGGCACTCCCGGTGGCATCTGCACTCTTGAAGACAGACTTGTG 540  
DB 481 GAGAGATTAAGTGGCACTCCCGGTGGCATCTGCACTCTTGAAGACAGACTTGTG 540  
QY 541 AATCTAGAGTGTGTAACATGAAATCTCAAGTGAAGCTTCTGAGATTCAGATTCAGAG 600  
DB 541 AATCTAGAGTGTGTAACATGAAATCTCAAGTGAAGCTTCTGAGATTCAGATTCAGAG 600  
QY 601 AGAAGAGATTTGAGGTTTGAAGCAAAATGACGTGAACAGCAACCAATCCAATGTGTA 660  
DB 601 AGAAGAGATTTGAGGTTTGAAGCAAAATGACGTGAACAGCAACCAATCCAATGTGTA 660  
QY 661 ATTGAAGCTTTGAGTCTCACTTACCCGTTGGTAACCCCACTCTCAAGCTCTCTG 720  
DB 661 ATTGAAGCTTTGAGTCTCACTTACCCGTTGGTAACCCCACTCTCAAGCTCTCTG 720  
QY 721 AATATTCCTGTTTACCCCGCAAAATATTTACAGTATCTGAGAGAGTCTTTGGCCAG 780  
DB 721 AATATTCCTGTTTACCCCGCAAAATATTTACAGTATCTGAGAGAGTCTTTGGCCAG 780  
QY 781 GAGAAAGCAAGATCTGATCTTCAAGCAGATCCAGATTTTCAAGTCCCAATTTCAAG 840  
DB 781 GAGAAAGCAAGATCTGATCTTCAAGCAGATCCAGATTTTCAAGTCCCAATTTCAAG 840



OY	841	GCAGTTCGAACCTTACGACGAATGATGGCCATAAAACA	CTGCGCTGGATGAAATTCGACATT	900
Dp	841	GCAGTTCGAACCTTACGACGAATGATGGCCATAAAACA	CTGCGCTGGATGAAATTCGACATT	900
OY	901	TCAAAATACAGACTTTTCTCTATCAGCCTGGAGATGCT	CACGCTGATCTGCGCTAACGAT	960
Dp	901	TCAAAATACAGACTTTTCTCTATCAGCCTGGAGATGCT	CACGCTGATCTGCGCTAACGAT	960
OY	961	GATTCCTGAGGTACAAAGCTTACCTCCAAAGACTGGAG	CTTGAAATGAAAGAGCACTGC	1020
Dp	961	GATTCCTGAGGTACAAAGCTTACCTCCAAAGACTGGAG	CTTGAAATGAAAGAGCACTGC	1020
OY	1021	GTCCCTTTTGAAATAATTAAGGCGACACAAAGAAAGAA	AGGAGCTTACCTTACCCGACATATA	1080
Dp	1021	GTCCCTTTTGAAATAATTAAGGCGACACAAAGAAAGAA	AGGAGCTTACCTTACCCGACATATA	1080
OY	1081	CCTGGGGGAGTGTCTCTCAGTCAATTTTATCCTGGTGT	CTTGAATCCGAGCAATTCCT	1140
Dp	1081	CCTGGGGGAGTGTCTCTCAGTCAATTTTATCCTGGTGT	CTTGAATCCGAGCAATTCCT	1140
OY	1141	AAAAAGGCAITTTTGCAGAGCCCTGTGTGACTATAC	CACTGACAGTGTGAAAGCGCAGG	1200
Dp	1141	AAAAAGGCAITTTTGCAGAGCCCTGTGTGACTATAC	CACTGACAGTGTGAAAGCGCAGG	1200
OY	1201	CTACAGGAGCTGTGTGACTATTAACAGGGGACGCCAT	TATACCGCTTTGTATCAGATATCC	1260
Dp	1201	CTACAGGAGCTGTGTGACTATTAACAGGGGACGCCAT	TATACCGCTTTGTATCAGATATCC	1260
OY	1261	TGTGCTGTCTGTTGGATCTCTCTCTGCTTCCCTTCT	GTGCGAGGCACTGACATGCTC	1320
Dp	1261	TGTGCTGTCTGTTGGATCTCTCTCTGCTTCCCTTCT	GTGCGAGGCACTGACATGCTC	1320
OY	1321	CTGCTCGAACAACCTTCTCTAACTTCAACCCAGACCAT	ATTGCTGTGCAAGCTCAAGTTTA	1380
Dp	1321	CTGCTCGAACAACCTTCTCTAACTTCAACCCAGACCAT	ATTGCTGTGCAAGCTCAAGTTTA	1380
OY	1381	TTTTCACCCAGAAAGCTCCATTTTGTCTTCAACAT	TGTGTGAATTTCTGTCTATCGCAC	1440
Dp	1381	TTTTCACCCAGAAAGCTCCATTTTGTCTTCAACAT	TGTGTGAATTTCTGTCTATCGCAC	1440
OY	1441	ACAGAGGTTCTGCGGGAAGGAGATATGATCAGGCTGT	GCGCTTGTGGTTGCTTCACTT	1500
Dp	1441	ACAGAGGTTCTGCGGGAAGGAGATATGATCAGGCTGT	GCGCTTGTGGTTGCTTCACTT	1500
OY	1501	CTTCAGCCAAACATACATGATCCCATGAAGACAGCGG	AAAGCCCTGAGCTCTTAAGATA	1560
Dp	1501	CTTCAGCCAAACATACATGATCCCATGAAGACAGCGG	AAAGCCCTGAGCTCTTAAGATA	1560
OY	1561	TCCATCTCTCTCGAACAACAATTCCTTCCACTTAC	CCAGATGACCCCTCAATCCCATC	1620
Dp	1561	TCCATCTCTCTCGAACAACAATTCCTTCCACTTAC	CCAGATGACCCCTCAATCCCATC	1620
OY	1621	ATATATGTGTGGTCCAGGAACCGGCAATGACCCCGTT	ATATGGGTTCTTACACATATAGAG	1680
Dp	1621	ATATATGTGTGGTCCAGGAACCGGCAATGACCCCGTT	ATATGGGTTCTTACACATATAGAG	1680
OY	1681	AAACTCCAAAGAACACCCAGATGGAATTTTGGAGCAT	ATGTGGTTGTTTTTGGCTGC	1740
Dp	1681	AAACTCCAAAGAACACCCAGATGGAATTTTGGAGCAT	ATGTGGTTGTTTTTGGCTGC	1740
OY	1741	AGGCAATAAGATAGGATATCTATTCGAAAAGAGCT	CAGACATTCCTTAAGCATGGG	1800
Dp	1741	AGGCAATAAGATAGGATATCTATTCGAAAAGAGCT	CAGACATTCCTTAAGCATGGG	1800
OY	1801	ATCTTAACTCATCTTAAAGGTTTCTTCTTCAAGAGAT	GTCTCTGTGTGGGAGAGAGAACCC	1860
Dp	1798	ATCTTAACTCATCTTAAAGGTTTCTTCTTCAAGAGAT	GTCTCTGTGTGGGAGAGAGAACCC	1857
OY	1861	CCAGCAAAATATGTACAGAACACATCCAGCTTCAAT	GTGCGCAGGAGTGGCGAGATCTCTC	1920
Dp	1858	CCAGCAAAATATGTACAGAACACATCCAGCTTCAAT	GTGCGCAGGAGTGGCGAGATCTCTC	1917
OY	1921	CTCCAGAGAAACGCGCATATTTATGTGTGTGGAAGAT	GCAAGATATGCGCAAGATGTATA	1980

Accession	Sequence	Year
Db	CTCCAGAGAAAGGCCCATTTATATGTGTGTGGAGATCCAAAGAAATATGGCCCAAGAGATGA	1977
Qy	CATGATGCCCTTGTGCAATPATATAGCAAAAGAGTTGGAGTTGAAAACTAGAAAGCATG	2040
Db	CATGATGCCCTTGTGCAAAATPATATAGCAAAAGAGTTGGAGTTGAAAACTAGAAAGCATG	2037
Qy	AAAACCCGCGCACTTTAAAGAAGAAAAAGCTACCTTCAGATATTTGGTCATA	2097
Db	AAAACCCGCGCACTTTAAAGAAGAAAAAGCTACCTTCAGATATTTGGTCATA	2094

```

? RESULT 6
? US-09-371-347-47
? Sequence 47, Application US/09371347
? Publication NO. US20030082676A1
? GENERAL INFORMATION:
? APPLICANT: Roy A. Gravel et al.
? TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE REDUCTASE:
? TITLE OF INVENTION: CLONING, AND METHODS FOR EVALUATING RISK OF NEURAL TUBE
? TITLE OF INVENTION: DEFECTS, CARDIOVASCULAR DISEASE, AND CANCER
? FILE REFERENCE: 50004/003003
? CURRENT APPLICATION NUMBER: US/09/371,347
? CURRENT FILING DATE: 1999-08-10
? PRIOR APPLICATION NUMBER: 60/071,622
? PRIOR FILING DATE: 1998-01-16
? PRIOR APPLICATION NUMBER: 09/232,028
? PRIOR FILING DATE: 1999-01-15
? NUMBER OF SEQ ID NOS: 51
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 47
? LENGTH: 2093
? TYPE: DNA
? ORGANISM: Homo sapiens
? US-09-371-347-47

```

Query Match	Best Local Match	Similarity	83.1%;	Score 1742;	DB 10;	Length 2093;
Matches 2092;	Conservative	0;	Mismatches	1;	Indels	4;
						Gaps 1;
Qy	1	ATGAGGAGGTTCTGTTACTATATGTCACACAGCGGAGCAGCGAAAGGCCATCGCAGAA	60			
Db	1	ATGAGGAGGTTCTGTTACTATATGTCACACAGCGGAGCAGCGAAAGGCCATCGCAGAA	60			
Qy	61	GAAATATGTGAGCAAGCTGTGTATCATGTATTTCTGCAATCTTCACTGTATTAGTGA	120			
Db	61	GAAATGTGTGAGCAAGCTGTGTATCATGTATTTCTGCAAGTCTTCACTGTATTAGTGA	120			
Qy	121	TCCGATTAAGTATGACTTAAACCGGAAACAGCTCCTTGTGTGTGTGTGAGTTTCTAACACG	180			
Db	121	TCCGATTAAGTATGACTTAAACCGGAAACAGCTCCTTGTGTGTGTGTGAGTTTCTAACACG	180			
Qy	181	GGCACCGGAGACCCAGCCGACACAGCCCGCAAGTTTGTTAAGAAATACAGAACCAACA	240			
Db	181	GGCACCGGAGACCCAGCCGACACAGCCCGCAAGTTTGTTAAGAAATACAGAACCAACA	240			
Qy	241	CTGC CGGTGATTTCTTTGCTCACTCGCGTATGGGTTACTGGTCTCGGTGATTCAGAA	300			
Db	241	CTGC CGGTGATTTCTTTGCTCACTCGCGTATGGGTTACTGGTCTCGGTGATTCAGAA	300			
Qy	301	TACACTTACTTTTGCAATGGGGGGAAGTAATTATTAACGACTTCAAGAGCTTGAAGCC	360			
Db	301	TACACTTACTTTTGCAATGGGGGGAAGTAATTATTAACGACTTCAAGAGCTTGAAGCC	360			
Qy	361	CGGCAATTCATATGACATGACATGACAGATGACTGTGTAGGTTTGAACCTTGTGTGAG	420			
Db	361	CGGCAATTCATATGACATGACATGACAGATGACTGTGTAGGTTTGAACCTTGTGTGAG	420			
Qy	421	CCGTGATATGCTGTGACCTCTGGCCAGACCTTCAGAAAGCAATTTTAGTCAAGCAGACAGCAA	480			
Db	421	CCGTGATATGCTGTGACCTCTGGCCAGACCTTCAGAAAGCAATTTTAGTCAAGCAGACAGCAA	480			
Qy	481	GAGGAGATTAAGTGGCGCACTCCCGGTGGCATCACTGTGATCTTTGAGACAGACTTTGTG	540			

Db 481 GAGAGAGTAAGTGGCCGACCTCCGGTGGCATCACCTGCATCTTGGAGAGACGCTTGTG 540  
Qy 541 AAGTCAGAGCTGCTACATCACTTGAATCTCAAGTCAGAGCTTGCAGATTGCATGATTCAGGA 600  
Db 541 AAGTCAGAGCTGCTACATCACTTGAATCTCAAGTCAGAGCTTGCAGATTGCATGATTCAGGA 600  
Qy 601 AGAAGAGATTCTGAGGTTTGAAGCAAAATGCACTGAGCAAGCAACCAATCCAAATGTTGTA 660  
Db 601 AGAAGAGATTCTGAGGTTTGAAGCAAAATGCACTGAGCAAGCAACCAATCCAAATGTTGTA 660  
Qy 661 ATTGAAGACTTTAGTCTCCTACCTTACCCGTTGGTACCCCACTCTCAAGGCTCTCTG 720  
Db 661 ATTGAAGACTTTAGTCTCCTACCTTACCCGTTGGTACCCCACTCTCTCAAGGCTCTCTG 720  
Qy 721 AATATTCCTGTTTACCCCAAGATATTTACAGGTATCATGTCAGAGAGTCTTGGCCAG 780  
Db 721 AATATTCCTGTTTACCCCAAGATATTTACAGGTATCATGTCAGAGAGTCTTGGCCAG 780  
Qy 781 GAGGAAAGCCAAAGTATCTGTGACTTCAGCAGATCCAGTTTTCAGTCCAAATTTCAAG 840  
Db 781 GAGGAAAGCCAAAGTATCTGTGACTTCAGCAGATCCAGTTTTCAGTCCAAATTTCAAG 840  
Qy 841 GCAGTTCACTTACTAGCAATGATGCAATAAACCACTGCTGGTGAATGGACATT 900  
Db 841 GCAGTTCACTTACTAGCAATGATGCAATAAACCACTGCTGGTGAATGGACATT 900  
Qy 901 TCAATATCAGACTTTTCTATTCAGCTGAGAGTCCCTTCAAGCTGATCTCCCTAACAT 960  
Db 901 TCAATATCAGACTTTTCTATTCAGCTGAGAGTCCCTTCAAGCTGATCTCCCTAACAT 960  
Qy 961 GATTCTGAGGTACCAAGCTTACTCCAAAGACTGCAAGCTTGAAGTAAAGAGCACTGC 1020  
Db 961 GATTCTGAGGTACCAAGCTTACTCCAAAGACTGCAAGCTTGAAGTAAAGAGCACTGC 1020  
Qy 1021 GTGCTTTGAAATTAAGAGGACACAAAGAAAGAGACTCCCTTACCCGCAATATA 1080  
Db 1021 GTGCTTTGAAATTAAGAGGACACAAAGAAAGAGACTCCCTTACCCGCAATATA 1080  
Qy 1081 CCTGCGGAGATGTTCTCTCAAGTTCATTTTACCTGTGTCTTGAATCCGAGCAATTCCT 1140  
Db 1081 CCTGCGGAGATGTTCTCTCAAGTTCATTTTACCTGTGTGTGAATCCGAGCAATTCCT 1140  
Qy 1141 AAAAAGGCAATTTTGGCAGCCCTTGTGACTATTCAGTCAAGTGTGAAAAGCCGAG 1200  
Db 1141 AAAAAGGCAATTTTGGCAGCCCTTGTGACTATTCAGTCAAGTGTGAAAAGCCGAG 1200  
Qy 1201 CTACAGAGCTGTGCACTAAACAAAGGGGCAAGCGATATATAGCCGCTTGTACAGATGCC 1260  
Db 1201 CTACAGAGCTGTGCACTAAACAAAGGGGCAAGCGATATATAGCCGCTTGTACAGATGCC 1260  
Qy 1261 TGTGCTGCTTGTGTGATCTCTCTGCTTCCCTTCTTGCAGACCACTCAAGTCTC 1320  
Db 1261 TGTGCTGCTTGTGTGATCTCTCTGCTTCCCTTCTTGCAGACCACTCAAGTCTC 1320  
Qy 1321 CTGCTCGAATCTTCTCTAACTTCAACCCAGACCAATATTCGTGTGCAAGCTCAAGTTA 1380  
Db 1321 CTGCTCGAATCTTCTCTAACTTCAACCCAGACCAATATTCGTGTGCAAGCTCAAGTTA 1380  
Qy 1381 TTTCAACCCAGAAAGCTCAATTTTGTCTCAACATGTGAAATTTCTGTCTACGCGACA 1440  
Db 1381 TTTCAACCCAGAAAGCTCAATTTTGTCTCAACATGTGAAATTTCTGTCTACGCGACA 1440  
Qy 1441 ACAGAGTTCTGCGAAGGAGATATGTAAGCTGCTGCTGTGTGTTGCTTCAAGT 1500  
Db 1441 ACAGAGTTCTGCGAAGGAGATATGTAAGCTGCTGCTGTGTGTTGCTTCAAGT 1500  
Qy 1501 CTTCAGGCAAAACATCATGCTCATGAAAGACGCGGAAAGCCCTGCTCTTAAGATA 1560  
Db 1501 CTTCAGGCAAAACATCATGCTCATGAAAGACGCGGAAAGCCCTGCTCTTAAGATA 1560  
Qy 1561 TCCATCTCTCTCGAACAACAATTTCTTCACTTACAGATGACCCCTCAATCCCCTATC 1620

Db 1561 TCCATCTCTCTCGAACAACAATTTCTTCACTTACAGATGACCCCTCAATCCCCTATC 1620  
Qy 1621 ATATAGTGGGTCCAGAAACCGGACATAGCCCGTTATTGGGTCTCTACAACTAGAGAG 1680  
Db 1621 ATATAGTGGGTCCAGAAACCGGACATAGCCCGTTATTGGGTCTCTACAACTAGAGAG 1680  
Qy 1681 AAATCTCAAGAACACACCCAGATGAAATTTTGGAGCAATGTGTTGTTTGGCTGC 1740  
Db 1681 AAATCTCAAGAACACACCCAGATGAAATTTTGGAGCAATGTGTTGTTTGGCTGC 1740  
Qy 1741 AGGCATTAAGATGAGGATTAATCTATTCAGAAAGAGCTCAGACATTTCTTAAGCATGG 1800  
Db 1741 AGGCATTAAGATGAGGATTAATCTATTCAGAAAGAGCTCAGACATTTCTTAAGCATGG 1800  
Qy 1797 ATCTTAATCATCTTAAGGTTTCTTCTCAAGAGATGCTCTGTGGAGAGAGAAAGCC 1856  
Db 1797 ATCTTAATCATCTTAAGGTTTCTTCTCAAGAGATGCTCTGTGGAGAGAGAAAGCC 1856  
Qy 1861 CCAGCAAGTATGTACAGACAAATCCAGCTTCATGCGCAGAGGTGCGAATTCCTC 1920  
Db 1861 CCAGCAAGTATGTACAGACAAATCCAGCTTCATGCGCAGAGGTGCGAATTCCTC 1920  
Qy 1921 CTCAGAGAGAGGCGCATTTTATGTGTGAGATGCAAAAGATATGCGCAAGATGTA 1980  
Db 1921 CTCAGAGAGAGGCGCATTTTATGTGTGAGATGCAAAAGATATGCGCAAGATGTA 1980  
Qy 1981 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGTTGAGTTGAAAATCTAGAGCAATG 2040  
Db 1981 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGTTGAGTTGAAAATCTAGAGCAATG 2040  
Qy 1977 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGTTGAGTTGAAAATCTAGAGCAATG 2036  
Db 1977 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGTTGAGTTGAAAATCTAGAGCAATG 2036  
Qy 2041 AAAACCTTGGCCCTTTAAAGAAAGAAACGCTTACCTTCAGATATTTGGTCAATA 2097  
Db 2041 AAAACCTTGGCCCTTTAAAGAAAGAAACGCTTACCTTCAGATATTTGGTCAATA 2097  
  
RESULT 7  
US-10-741-600-692  
; Sequence 692, Application US/10741600  
; Publication No. US20050026169A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; FILE REFERENCE: CL001499  
; CURRENT APPLICATION NUMBER: US/10/741,600  
; CURRENT FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 73997  
; SOFTWARE: PASTSEQ for Windows Version 4.0  
; SEQ ID NO 692  
; LENGTH: 3256  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-741-600-692  
  
Query Match 50.6%; Score 1062; DB 21; Length 3256;  
Best Local Similarity 99.1%; Pred. No. 0;  
Matches 2012; Conservative 0; Mismatches 19; Indels 0; Gaps 0;  
  
Qy 67 TGTGAGCAAGCTGTGTAATGATGATTTTCTGAGATCTTCACTGATTAATGATCCGAT 126  
Db 160 TGTGAGCAAGCTGTGTAATGATGATTTTCTGAGATCTTCACTGATTAATGATCCGAT 219  
Qy 127 AAGTATGACTTAATAAAGCGAAACAGCTCTCTTGTGTGTGTGTTTCTACAGGCGACC 186  
Db 220 AAGTATGACTTAATAAAGCGAAACAGCTCTCTTGTGTGTGTGTTTCTACAGGCGACC 279  
Qy 187 GGAAGCCCAACCGGACAGCCGCAAGTTGTTTAAAGAAATACAGAAACCAACCTGCGC 246  
Db 280 GGAAGCCCAACCGGACAGCCGCAAGTTGTTTAAAGAAATACAGAAACCAACCTGCGC 339  
Qy 247 GTTATTTCTTGTGCTACCGGTATGAGTTTCTGAGGTCTCGGTGTTTCAAGATTCACC 306  
Db 340 GTTATTTCTTGTGCTACCGGTATGAGTTTCTGAGGTCTCGGTGTTTCAAGATTCACC 399

QY 307 TACTTTGCAATGAGGGAAGATTAATTGATAAAGCACTTCAAGAGCTTGAAGCCGGCAT 366  
 DB 400 TACTTTGCAATGAGGGAAGATTAATTGATAAAGCACTTCAAGAGCTTGAAGCCGGCAT 459  
 QY 367 TTTCTATGACATGACATGACATGACATGACATGACATGACATGACATGACATGACAT 426  
 DB 460 TTTCTATGACATGACATGACATGACATGACATGACATGACATGACATGACATGACAT 519  
 QY 427 ATTGCTGACATGACATGACATGACATGACATGACATGACATGACATGACATGACAT 486  
 DB 520 ATTGCTGACATGACATGACATGACATGACATGACATGACATGACATGACATGACAT 579  
 QY 487 ATTAAGTGGGCACTCCGGTGGGCACTGACATGACATGACATGACATGACATGACAT 546  
 DB 580 ATTAAGTGGGCACTCCGGTGGGCACTGACATGACATGACATGACATGACATGACAT 639  
 QY 547 GAGCTGCTACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 606  
 DB 640 GAGCTGCTACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 699  
 QY 607 GATTCGAGCTTTTGAAGCAAAATGACATGACATGACATGACATGACATGACATGACAT 666  
 DB 700 GATTCGAGCTTTTGAAGCAAAATGACATGACATGACATGACATGACATGACATGACAT 759  
 QY 667 GACTTGAAGCTCTCACTTACCCGTTGGGTAACCCGCACTCTCAAGAGCTCTGAAATTT 726  
 DB 760 GACTTGAAGCTCTCACTTACCCGTTGGGTAACCCGCACTCTCAAGAGCTCTGAAATTT 819  
 QY 727 CCTGCTTTACCCGCAAAATTTTACAGATACATCTGACAGAGTCTCTTGGCCAGAGAA 786  
 DB 820 CCTGCTTTACCCGCAAAATTTTACAGATACATCTGACAGAGTCTCTTGGCCAGAGAA 879  
 QY 787 AGCCAGATCTGATGATCTGACATGACATGACATGACATGACATGACATGACATGACAT 846  
 DB 880 AGCCAGATCTGATGATCTGACATGACATGACATGACATGACATGACATGACATGACAT 939  
 QY 847 CAATCTACTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 906  
 DB 940 CAATCTACTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 999  
 QY 907 ACAGACTTTTCTATGACATGACATGACATGACATGACATGACATGACATGACATGACAT 966  
 DB 1000 ACAGACTTTTCTATGACATGACATGACATGACATGACATGACATGACATGACATGACAT 1059  
 QY 967 GAGGTACAAAGCTTCTCAAGAGCTGACATGACATGACATGACATGACATGACATGACAT 1026  
 DB 1060 GAGGTACAAAGCTTCTCAAGAGCTGACATGACATGACATGACATGACATGACATGACAT 1119  
 QY 1027 TTGAAAATTAAGGCAACAAGAAAGAAAGGATACCTTACCCAGATTAATACCTGAG 1086  
 DB 1120 TTGAAAATTAAGGCAACAAGAAAGAAAGGATACCTTACCCAGATTAATACCTGAG 1179  
 QY 1087 GATGATCTCTCAAGTCAATTTTACCTGATCTTGAATCCGAGCAATCTCTTAAAG 1146  
 DB 1180 GATGATCTCTCAAGTCAATTTTACCTGATCTTGAATCCGAGCAATCTCTTAAAG 1239  
 QY 1147 GCATTTTGGAGCCCTTGTGACATTAACAGTGAAGTGTGAAAAGCGAGGCTTACAG 1206  
 DB 1240 GCATTTTGGAGCCCTTGTGACATTAACAGTGAAGTGTGAAAAGCGAGGCTTACAG 1299  
 QY 1207 GAGCTGTGACATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1266  
 DB 1300 GAGCTGTGACATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1359  
 QY 1267 TGGTTTGGATCT 1326  
 DB 1360 TGGTTTGGATCT 1419  
 QY 1327 GAAATCTTCTTAACTTCAACCAAGCAATATGCTGACAGCTCAAGTATTTTAC 1386  
 DB 1420 GAAATCTTCTTAACTTCAACCAAGCAATATGCTGACAGCTCAAGTATTTTAC 1479

QY 1387 CCAGAAAGCTCATTCTTCTTCAACATTTGGAATTTCTGTCTATCTGCCAACAAGAG 1446  
 DB 1480 CCAGAAAGCTCATTCTTCTTCAACATTTGGAATTTCTGTCTATCTGCCAACAAGAG 1539  
 QY 1447 GTTCTGCGAAGGAGATATGATCAGGCTGGCTGGCTTGTGTGCTTCAAGTCTTTCAG 1506  
 DB 1540 GTTCTGCGAAGGAGATATGATCAGGCTGGCTGGCTTGTGTGCTTCAAGTCTTTCAG 1599  
 QY 1507 CCAACATACATGATCCCATGAAGACAGGAGGAAAGCCCTGAGCTCTAAGATATCATC 1566  
 DB 1600 CCAACATACATGATCCCATGAAGACAGGAGGAAAGCCCTGAGCTCTAAGATATCATC 1659  
 QY 1567 TCTCTCGAACAACAATTTCTTCACTTACAGATGACATCCCTCAATCCCATATATG 1626  
 DB 1660 TCTCTCGAACAACAATTTCTTCACTTACAGATGACATCCCTCAATCCCATATATG 1719  
 QY 1627 GTGGGTCCAGAAACCGGATAGCCCGCTTATTTGGGTTCTTACATATGAGAAATCTC 1686  
 DB 1720 GTGGGTCCAGAAACCGGATAGCCCGCTTATTTGGGTTCTTACATATGAGAAATCTC 1779  
 QY 1687 CAAGAACACACCGATGAAATTTTGGAGCAATGTTGTTTGGCTGAGGAT 1746  
 DB 1780 CAAGAACACACCGATGAAATTTTGGAGCAATGTTGTTTGGCTGAGGAT 1839  
 QY 1747 AAGATAGGATATATCTATTCAGAAAGAGCTCAGACATTTCTTAAAGATGAGATCTTA 1806  
 DB 1840 AAGATAGGATATATCTATTCAGAAAGAGCTCAGACATTTCTTAAAGATGAGATCTTA 1899  
 QY 1807 ACTCATTAAGGTTTCTTCTTCAAGAGATGCTCTGTTGGGAGAGAGAGCCCAAGA 1866  
 DB 1900 ACTCATTAAGGTTTCTTCTTCAAGAGATGCTCTGTTGGGAGAGAGAGCCCAAGA 1959  
 QY 1867 AAGTATGTCAGAACACATCCAGATTCATGAGCCAGAGTGGCGAATCTCTCCAG 1926  
 DB 1960 AAGTATGTCAGAACACATCCAGATTCATGAGCCAGAGTGGCGAATCTCTCCAG 2019  
 QY 1927 GAGAACGCGCATATTTATGTTGTGAGATGACAAAGATATGAGATGATACATGAT 1986  
 DB 2020 GAGAACGCGCATATTTATGTTGTGAGATGACAAAGATATGAGATGATACATGAT 2079  
 QY 2047 CTGGCCACTTTTAAAGAAAGAAAGCGCTACCTTCAAGATTTTGTGATCA 2097  
 DB 2140 CTGGCCACTTTTAAAGAAAGAAAGCGCTACCTTCAAGATTTTGTGATCA 2190

## RESULT 8

US-10-741-600-693  
 ; Sequence 693, Application US/10741600  
 ; Publication No. US20050026169A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CARGILL, Michele et al.  
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
 ; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CL001499  
 ; CURRENT APPLICATION NUMBER: US/10/741,600  
 ; NUMBER OF SEQ ID NOS: 73997  
 ; SOFTWARE: PasteSeq for Windows Version 4.0  
 ; SEQ ID NO 693  
 ; LENGTH: 3274  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-741-600-693

Query Match 50.6%; Score 1062; DB 21; Length 3274;  
 Best Local Similarity 99.1%; Pred. No. 0;  
 Matches 2012; Conservative 0; Mismatches 19; Indels 0; Gaps 0;  
 QY 67 TGTGAGCAAGCTGTGTGATCATGATTTTCTGCAAGATCTTCACTGTATTAATCCGAT 126

178 TGAGAGACGCTGGGACATGATGATTTTCGACAGATCTTCAGTATTAATGATCCGAT 237  
127 AAGTATGACCTAAACCCGAAAGAGCTCTCTGTGTGTGTGTGTCTACACGGCACC 186  
238 AAGTATGACCTAAACCCGAAAGAGCTCTCTGTGTGTGTGTGTCTACACGGCACC 297  
187 GGAAGACCCACCCGACACAGCCCGGAGTTGTTAAGAAATACAAACCAACATCTGCC 246  
298 GGAAGACCCACCCGACACAGCCCGGAGTTGTTAAGAAATACAAACCAACATCTGCC 357  
247 GTGATTTCTTGTCTACCTGGGTAAGGTTACTGGGTCCTGGTATTCAGAAATCACCC 306  
358 GTTGAATTTCTTGTCTACCTGGGTAAGGTTACTGGGTCCTGGTATTCAGAAATCACCC 417  
307 TACTTTTGAAGTGGGGGAGATTAATGATTAACGACTTCAGAGCTTTGAGCCCGCAT 366  
418 TACTTTTGAAGTGGGGGAGATTAATGATTAACGACTTCAGAGCTTTGAGCCCGCAT 477  
367 TTTCTATGACACTGGAACATGAGATGACTGTGATGTTAAGACTTGTGTGAGCCGTGG 426  
478 TTTCTATGACACTGGAACATGAGATGACTGTGATGTTAAGACTTGTGTGAGCCGTGG 537  
427 ATTGCTGGAATCTGGCCAGCCCTCAGAAAGCATTTTATGATCAAGCAGAGGACAAAGGAG 486  
538 ATTGCTGGAATCTGGCCAGCCCTCAGAAAGCATTTTATGATCAAGCAGAGGACAAAGGAG 597  
487 ATAAAGTGGCAGCTCCCGGTGGATCACTGCACTCTTGAAGACAGACCTTGTGAAGTCA 546  
598 ATAAAGTGGCAGCTCCCGGTGGATCACTGCACTCTTGAAGACAGACCTTGTGAAGTCA 657  
547 GAGCTGCTACACATTTGATCTCAAGTGTGAGCTTGTGATTCAGATGATTCAGAAAGAG 606  
658 GAGCTGCTACACATTTGATCTCAAGTGTGAGCTTGTGATTCAGATGATTCAGAAAGAG 717  
607 GATTCTGAGGTTTGAAGAAATGAGTGAACGAAACCAATGTTGTAATTGAA 666  
718 GATTCTGAGGTTTGAAGAAATGAGTGAACGAAACCAATGTTGTAATTGAA 777  
667 GACTTTGAGTCTCACTTACCCGTTCCGGTACCCCACTCTCAAGAGCTCTCTGAATATT 726  
778 GACTTTGAGTCTCACTTACCCGTTCCGGTACCCCACTCTCAAGAGCTCTCTGAATATT 837  
727 CCGTGTTCACCCCGAGAAATTTTACAGTACATCTGCAAGAGTCTTGGCCAGAGGAA 786  
838 CCGTGTTCACCCCGAGAAATTTTACAGTACATCTGCAAGAGTCTTGGCCAGAGGAA 897  
787 AGCCAAAGTACTGTGACTTCAGAGAGATCCAGTTTTCAGATGCCAATTTGAAAGGAGTT 846  
898 AGCCAAAGTACTGTGACTTCAGAGAGATCCAGTTTTCAGATGCCAATTTGAAAGGAGTT 957  
847 CAATTACTACGATGATGCAATTAACCACTCTGCTGTGATGAAATGACATTTCAAT 906  
958 CAATTACTACGATGATGCAATTAACCACTCTGCTGTGATGAAATGACATTTCAAT 1017  
907 ACAGACTTTTCTCTACGCTCGAGATGCTTCAAGCTGATCTGCCCTTAACAGTATCT 966  
1018 ACAGACTTTTCTCTACGCTCGAGATGCTTCAAGCTGATCTGCCCTTAACAGTATCT 1077  
967 GAGGTACAAAGCCCTACCTCAAGAGCTGAGCTTGAAGATTAAGAGAGCACTGGTCTT 1026  
1078 GAGGTACAAAGCCCTACCTCAAGAGCTGAGCTTGAAGATTAAGAGAGCACTGGTCTT 1137  
1027 TTGAAATTAAGGACGACACAAAGAGAAAGAGACTTACCTTACCCAGCATATACCTGCG 1086  
1138 TTGAAATTAAGGACGACACAAAGAGAAAGAGACTTACCTTACCCAGCATATACCTGCG 1197  
1087 GGAATGTTCTCTCCAGTTCATTTTACTGCTGTGTGAAATCCGAGCAATTCCTAAAAAG 1146  
1198 GGAATGTTCTCTCCAGTTCATTTTACTGCTGTGTGAAATCCGAGCAATTCCTAAAAAG 1257  
1147 GCATTTTGGAGGCTTGTGACTATACAGAGACAGTCTGAAAGCCCAAGCTACAG 1206

1258 GCATTTTTCGAGCCCTTGTGACTATACAGTACAGTGTGTGAAAGCCGAGGCTACAG 1317  
1207 GAGCTGTGACGTAAACAAAGGGGACGCCGATTTATAGCCGCTTGTGTACAGATGCTGTGCC 1266  
1318 GAGCTGTGACGTAAACAAAGGGGACGCCGATTTATAGCTGTGTGTGTGTGTGTGTGTGCC 1377  
1267 TGT 1326  
1378 TGT 1437  
1327 GAAACATCTTCTTAAACCTTCAACCCAGACCAATATGCTGTGACAGTCAAGTTATTTTAC 1386  
1438 GAAACATCTTCTTAAACCTTCAACCCAGACCAATATGCTGTGACAGTCAAGTTATTTTAC 1497  
1387 CCAGAAAGCTCCATTTTGTCTTCAACATTTGTGAAATTTCTGTCTACCTGCAACAGAG 1446  
1498 CCAGAAAGCTCCATTTTGTCTTCAACATTTGTGAAATTTCTGTCTACCTGCAACAGAG 1557  
1447 GTTCTGCGGAAAGGAGATATGATACAGCTGTGCTGTGCTGTGTGTGTGTGTGTGTGTGT 1506  
1558 GTTCTGCGGAAAGGAGATATGATACAGCTGTGCTGTGCTGTGTGTGTGTGTGTGTGTGT 1617  
1507 CCAGAAATATCATGATCCCATGAAAGACAGGGGAAAGCCCTGAGCTCTTAAGATATCATC 1566  
1618 CCAGAAATATCATGATCCCATGAAAGACAGGGGAAAGCCCTGAGCTCTTAAGATATCATC 1677  
1567 TCTCTCTGAAACCAAAATTTCTTCACTTACAGATGACCCCTCAATTCCTCATATATG 1626  
1678 TCTCTCTGAAACCAAAATTTCTTCACTTACAGATGACCCCTCAATTCCTCATATATG 1737  
1627 GTGGGTTCAGAAACCGGACATAGCCCGCTTATTTGGTTCCTTACAAATGAGAGAAATC 1686  
1738 GTGGGTTCAGAAACCGGACATAGCCCGCTTATTTGGTTCCTTACAAATGAGAGAAATC 1797  
1687 CAAGAACCAACCCAGATGAAATTTTGAAGCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1746  
1798 CAAGAACCAACCCAGATGAAATTTTGAAGCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1857  
1747 AAGGATAGGATTTATCTATTCAGAAAGAGCTCAGACATTTCTTAAAGATGGGATCTTA 1806  
1858 AAGGATAGGATTTATCTATTCAGAAAGAGCTCAGACATTTCTTAAAGATGGGATCTTA 1917  
1807 ACTCATCTAAAGTTTCTTCTCAAGAGATGCTCTGTGTGGGAGAGAGAGGAGGAGGAGGAG 1866  
1918 ACTCATCTAAAGTTTCTTCTCAAGAGATGCTCTGTGTGGGAGAGAGAGGAGGAGGAGGAG 1977  
1867 AAGTATGTACAAAGACCAACCTCAGCTTCAAGGACAGAGTGGCAGAAATCTCTCCAG 1926  
1978 AAGTATGTACAAAGACCAACCTCAGCTTCAAGGACAGAGTGGCAGAAATCTCTCCAG 2037  
1927 GAGAACGGCCATATTTATGTGTGTGAGATGCAAGAAATATGGCCAAAGATGTACATGAT 1986  
2038 GAGAACGGCCATATTTATGTGTGTGAGATGCAAGAAATATGGCCAAAGATGTACATGAT 2097  
1987 GCCCTGTGCAAAATTAAGCAAAAGAGTTGAGTTGAAAACTTAAAGCAATGAAAGAAC 2046  
2098 GCCCTGTGCAAAATTAAGCAAAAGAGTTGAGTTGAAAACTTAAAGCAATGAAAGAAC 2157  
2047 CTGGCCACTTTTAAAGAAAGAAAGAAAGCTACCTTCAAGATATTTGTGATAT 2097  
2158 CTGGCCACTTTTAAAGAAAGAAAGAAAGCTACCTTCAAGATATTTGTGATAT 2208

RESULT 9  
US-10-029-386-6369  
; Sequence 6369, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GE  
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO

FILE REFERENCE: AEMICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 34288  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 6369  
LENGTH: 591  
TYPE: DNA  
ORGANISM: Homo sapiens

OTHER INFORMATION: MAP TO AC008727.5  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.45  
OTHER INFORMATION: NT HIT: AF121205.1, EVALUE 0.00e+00  
OTHER INFORMATION: SWISSPROT HIT: Q9UPW0, EVALUE 3.30e+00  
OTHER INFORMATION: EST\_HUMAN HIT: AU132586.1, EVALUE 0.00e+00  
US-10-029-386-6369

Query Match 15.7%; Score 330; DB 16; Length 591;  
Best Local Similarity 99.7%; Pred. No. 1.2e-169;  
Matches 380; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 401 GTTTAGAACTTGTGTTGAGCCCTGATTTGCTGAGACTTGGCCCAAGCCCTCAGAAAGATT 460  
DB 38 GTTTAGAACTTGTGTTGAGCCCTGATTTGCTGAGACTTGGCCCAAGCCCTCAGAAAGATT 97  
QY 461 TTAGGTCAAGCAGAGCAAGAGAGATTAAGTGGCGCACTCCCGGATGATCAGCTGCAAT 520  
DB 98 TTAGGTCAAGCAGAGCAAGAGAGATTAAGTGGCGCACTCCCGGATGATCAGCTGCAAT 157  
QY 521 CCTTGAAGCAGAGCCTTGTGAAGTCAAGCTGCTACATTTGAATCTCAAGTGAAGCTTC 580  
DB 158 CCTGAGAGCAGAGCCTTGTGAAGTCAAGCTGCTACATTTGAATCTCAAGTGAAGCTTC 217  
QY 581 TGAGATTCATGATTTAGAGAAAGAGATTTCTGAGTTTGAAGCAAAATGCAAGTGAACA 640  
DB 218 TGAGATTCATGATTTAGAGAAAGAGATTTCTGAGTTTGAAGCAAAATGCAAGTGAACA 277  
QY 641 GCAACCAATCAATGTTGATTAATGAAGACTTTGAGTCTCACTTACCCGTTGGTACCCC 700  
DB 278 GCAACCAATCAATGTTGATTAATGAAGACTTTGAGTCTCACTTACCCGTTGGTACCCC 337  
QY 701 CACTCTCACAAGCCTCTGTAATATTCCTGTTTACCCCGAATATTTTACAGGTACATC 760  
DB 338 CACTCTCACAAGCCTCTGTAATATTCCTGTTTACCCCGAATATTTTACAGGTACATC 397  
QY 761 TGCAGAGTCTCTTGGCCAGG 781  
DB 398 TGCAGAGTCTCTTGGCCAGG 418

RESULT 10  
US-10-029-386-20100  
Sequence 20100, Application US/10029386  
Publication No. US20030194704A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
FILE REFERENCE: AEMICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 34288  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 20100  
LENGTH: 379  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC008727.5  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.45  
OTHER INFORMATION: NT HIT: g114729757, EVALUE 0.00e+00

OTHER INFORMATION: SWISSPROT HIT: Q9UPW0, EVALUE 1.80e+00  
OTHER INFORMATION: EST\_HUMAN HIT: AU132586.1, EVALUE 0.00e+00  
US-10-029-386-20100

Query Match 15.6%; Score 328; DB 16; Length 379;  
Best Local Similarity 99.7%; Pred. No. 1.4e-168;  
Matches 378; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 402 TTTAGAACTTGTGTTGAGCCCTGATTTGCTGAGACTTGGCCCAAGCCCTCAGAAAGATT 461  
DB 1 TTTAGAACTTGTGTTGAGCCCTGATTTGCTGAGACTTGGCCCAAGCCCTCAGAAAGATT 60  
QY 462 TAGGTCAAGCAGAGCAAGAGAGATTAAGTGGCGCACTCCCGGATGATCAGCTGCAAT 521  
DB 61 TAGGTCAAGCAGAGCAAGAGAGATTAAGTGGCGCACTCCCGGATGATCAGCTGCAAT 120  
QY 522 CTTGAGAGCAGAGCCTTGTGAAGTCAAGCTGCTACATTTGAATCTCAAGTGAAGCTTC 581  
DB 121 CTTGAGAGCAGAGCCTTGTGAAGTCAAGCTGCTACATTTGAATCTCAAGTGAAGCTTC 180  
QY 582 GAGATTCATGATTTAGAGAAAGAGATTTCTGAGTTTGAAGCAAAATGCAAGTGAACA 641  
DB 181 GAGATTCATGATTTAGAGAAAGAGATTTCTGAGTTTGAAGCAAAATGCAAGTGAACA 240  
QY 642 CAACCAATCAATGTTGATTAATGAAGACTTTGAGTCTCACTTACCCGTTGGTACCCC 701  
DB 241 CAACCAATCAATGTTGATTAATGAAGACTTTGAGTCTCACTTACCCGTTGGTACCCC 300  
QY 702 ACTCTCACAAGCCTCTGTAATATTCCTGTTTACCCCGAATATTTTACAGGTACATC 761  
DB 301 ACTCTCACAAGCCTCTGTAATATTCCTGTTTACCCCGAATATTTTACAGGTACATC 360  
QY 762 GCAGAGTCTCTTGGCCAG 780  
DB 361 GCAGAGTCTCTTGGCCAG 379

RESULT 11  
US-10-029-386-1735  
Sequence 1735, Application US/10029386  
Publication No. US20030194704A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
FILE REFERENCE: AEMICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 34288  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 1735  
LENGTH: 591  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC021609.3  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.4  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.8  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.2  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.2  
OTHER INFORMATION: NT HIT: AF121205.1, EVALUE 0.00e+00  
OTHER INFORMATION: EST\_HUMAN HIT: AU132586.1, EVALUE 0.00e+00  
OTHER INFORMATION: SWISSPROT HIT: Q9UPW0, EVALUE 3.30e+00  
US-10-029-386-1735

Query Match 13.3%; Score 279; DB 16; Length 591;  
Best Local Similarity 99.5%; Pred. No. 1.2e-141;  
Matches 379; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 401 GTTTAGAACTTGTGGTGAACCTGTGACTGTGCACTCTGGCCAGCCCTTCAGAAAGCATT 460  
DB 38 GTTTAGAACTTGTGGTGAACCTGTGACTGTGCACTCTGGCCAGCCCTTCAGAAAGCATT 97  
QY 461 TTAGGTCAGACAGAGCAAGAGAGATAGTGGCCGCACTCCGGTGGCATCACTGGCAT 520  
DB 98 TTAGGTCAGACAGAGCAAGAGAGATAGTGGCCGCACTCCGGTGGCATCACTGGCAT 157  
QY 521 CCTTGAAGACAGACCTTGTGAAGTCAGAGCTGCTACACATTGAATCTCAAGTCGAGCTTC 580  
DB 158 CCTGAGAGACAGACCTGTGAAGTCAGAGCTGCTACACATTGAATCTCAAGTCGAGCTTC 217  
QY 581 TGAGATTCAGATGATTCAGAAAGAAAGATTCAGAGTTTGAAGCAAAATTCAGTGAACA 640  
DB 218 TGAGATTCAGATGATTCAGAAAGAAAGATTCAGAGTTTGAAGCAAAATTCAGTGAACA 277  
QY 641 GCAACCAATCCAAATGTTGTAATGAAGACTTGTGAGCTGCTACACTTACCCGTTGGTACCCC 700  
DB 278 GCAACCAATCCAAATGTTGTAATGAAGACTTGTGAGCTGCTACACTTACCCGTTGGTACCCC 337  
QY 701 CACTTCACAAAGCCTCTGTAATATTCCTGTTTACCCCAAGATATTTACAGGTACATC 760  
DB 338 CACTTCACAAAGCCTCTGTAATATTCCTGTTTACCCCAAGATATTTACAGGTACATC 397  
QY 761 TGCAGAGTCTCTTGGCCAGG 781  
DB 398 TGCAGAGTCTCTTGGCCAGG 418

## RESULT 12

US-10-029-386-15435  
/ Sequence 15435, Application US/10029386  
/ Publication No. US20030194704A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Penn, Sharon G.  
/ APPLICANT: Hanzel, David R.  
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
/ FILE REFERENCE: AEMICA-X-2  
/ CURRENT APPLICATION NUMBER: US/10/029,386  
/ CURRENT FILING DATE: 2001-12-20  
/ NUMBER OF SEQ ID NOS: 34288  
/ SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
/ SEQ ID NO 15435  
/ LENGTH: 379  
/ TYPE: DNA  
/ ORGANISM: Homo sapiens  
/ FEATURE:  
/ OTHER INFORMATION: MAP TO AC021609.3  
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6  
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.4  
/ OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2  
/ OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.8  
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.2  
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.2  
/ OTHER INFORMATION: SWISSPROT HIT: Q9UPW0, EVALUE 1.80e+00  
/ OTHER INFORMATION: EST HUMAN HIT: AU132586.1, EVALUE 0.00e+00  
/ OTHER INFORMATION: NT HIT: gi14729757, EVALUE 0.00e+00  
US-10-029-386-15435

Query Match 13.2%; Score 277; DB 16; Length 379;  
Best Local Similarity 99.5%; Pred. No. 1.4e-140;

Matches 377; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 402 TTTTGAACCTGTGTGAGCCGTGAGTTGTGACTCTGGCCAGCCCTTCAGAAAGCATT 461  
DB 1 TTTTGAACCTGTGTGAGCCGTGAGTTGTGACTCTGGCCAGCCCTTCAGAAAGCATT 60  
QY 462 TAGGTCAAGAGAGAGCAAGAGAGATAGTGGCCATCCCGTGGCATCACTGGATC 521  
DB 61 TAGGTCAAGAGAGAGCAAGAGAGATAGTGGCCATCCCGTGGCATCACTGGATC 120

QY 522 CTTGAGAGACAGACCTTGTGAAGTCAGAGCTGCTACACATTGAATCTCAAGTCGAGCTTCT 581  
DB 121 CTCGAGAGACAGACCTTGTGAAGTCAGAGCTGCTACACATTGAATCTCAAGTCGAGCTTCT 180  
QY 582 GAGATTCGATGATTCAGAAAGAAAGATTCGAGGTTTGAAGCAAAATTCAGTGAACA 641  
DB 181 GAGATTCGATGATTCAGAAAGAAAGATTCGAGGTTTGAAGCAAAATTCAGTGAACA 240  
QY 642 CAACCAATCCAAATGTTGTAATGAAGACTTGTGAGCTGCTACACTTACCCGTTGGTACCCC 701  
DB 241 CAACCAATCCAAATGTTGTAATGAAGACTTGTGAGCTGCTACACTTACCCGTTGGTACCCC 300  
QY 702 ACTTCACAAAGCCTCTGTAATATTCCTGTTTACCCCAAGATATTTACAGGTACATC 761  
DB 301 ACTTCACAAAGCCTCTGTAATATTCCTGTTTACCCCAAGATATTTACAGGTACATC 360  
QY 762 GCAGAGTCTCTTGGCCAG 780  
DB 361 GCAGAGTCTCTTGGCCAG 379

## RESULT 13

US-10-741-600-17757  
/ Sequence 17757, Application US/10741600  
/ Publication No. US20050026169A1  
/ GENERAL INFORMATION:  
/ APPLICANT: CARGILL, Michele et al.  
/ TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
/ FILE REFERENCE: CL001499  
/ CURRENT APPLICATION NUMBER: US/10/741,600  
/ CURRENT FILING DATE: 2003-12-22  
/ NUMBER OF SEQ ID NOS: 73997  
/ SOFTWARE: FastSeq for Windows Version 4.0  
/ SEQ ID NO 17757  
/ LENGTH: 43985  
/ TYPE: DNA  
/ ORGANISM: Homo sapiens  
US-10-741-600-17757

Query Match 12.7%; Score 266; DB 21; Length 43985;  
Best Local Similarity 99.5%; Pred. No. 2e-134;

Matches 366; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 401 GTTTAGAACTTGTGTGAGCCGTGAGTTGTGACTCTGGCCAGCCCTTCAGAAAGCATT 460  
DB 14836 GTTTAGAACTTGTGTGAGCCGTGAGTTGTGACTCTGGCCAGCCCTTCAGAAAGCATT 14895  
QY 461 TTAGGTCAGACAGAGCAAGAGAGATAGTGGCCGCACTCCGGTGGCATCACTGGCAT 520  
DB 14896 TTAGGTCAGACAGAGCAAGAGAGATAGTGGCCGCACTCCGGTGGCATCACTGGCAT 14955  
QY 521 CCTTGAAGACAGACCTTGTGAAGTCAGAGCTGCTACACATTGAATCTCAAGTCGAGCTTC 580  
DB 14956 CCTTGAAGACAGACCTTGTGAAGTCAGAGCTGCTACACATTGAATCTCAAGTCGAGCTTC 15015  
QY 581 TGAGATTCAGATGATTCAGAAAGAAAGATTCGAGGTTTGAAGCAAAATTCAGTGAACA 640  
DB 15016 TGAGATTCAGATGATTCAGAAAGAAAGATTCGAGGTTTGAAGCAAAATTCAGTGAACA 15075  
QY 641 GCAACCAATCCAAATGTTGTAATGAAGACTTGTGAGCTGCTACACTTACCCGTTGGTACCCC 700  
DB 15076 GCAACCAATCCAAATGTTGTAATGAAGACTTGTGAGCTGCTACACTTACCCGTTGGTACCCC 15135  
QY 701 CACTTCACAAAGCCTCTGTAATATTCCTGTTTACCCCAAGATATTTACAGGTACATC 760  
DB 15136 CACTTCACAAAGCCTCTGTAATATTCCTGTTTACCCCAAGATATTTACAGGTACATC 15195  
QY 761 TGCAGAGG 788  
DB 15196 TGCAGAGG 15203



RESULT 14  
US-10-029-386-633/c  
; Sequence 633, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
; FILE REFERENCE: AEOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 633  
; LENGTH: 525  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC021609.3  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.48  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.58  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.52  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.57  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.79  
; OTHER INFORMATION: SWISSPROT HIT: P37039, EVALUOR 1.00e-06  
; OTHER INFORMATION: EST\_HUMAN HIT: BF346446.1, EVALUOR 1.00e-98  
; OTHER INFORMATION: NT HIT: AF121212.1, EVALUOR 0.00e+00  
US-10-029-386-633

Query Match 9.0%; Score 188; DB 16; Length 525;  
Best Local Similarity 100.0%; Pred. No. 1.1e-91;

Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1765 TTCAGAAAAGAGCTCAGACATTTCTTAAGCATGGATCTTAATCATCTAAAGTTTCC 1824  
DB 234 TTCAGAAAAGAGCTCAGACATTTCTTAAGCATGGATCTTAATCATCTAAAGTTTCC 175  
QY 1825 TTCTCAAGAGATGCTCTGTGGGAGAGAGAAAGCCCAAGAAAGTATGACAGCAAC 1884  
DB 174 TTCTCAAGAGATGCTCTGTGGGAGAGAGAAAGCCCAAGAAAGTATGACAGCAAC 115  
QY 1885 ATCCAGCTTCATGGCCAGACAGTGGGAGAAATCTCTCCAGAGAAAGGCCATATTAT 1944  
DB 114 ATCCAGCTTCATGGCCAGACAGTGGGAGAAATCTCTCCAGAGAAAGGCCATATTAT 55  
QY 1945 GTGTGTGG 1952  
DB 54 GTGTGTGG 47

RESULT 15

US-10-029-386-14338/c

; Sequence 14338, Application US/10029386  
; Publication No. US20030194704A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharon G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G

; FILE REFERENCE: AEOMICA-X-2

; CURRENT APPLICATION NUMBER: US/10/029,386

; CURRENT FILING DATE: 2001-12-20

; NUMBER OF SEQ ID NOS: 34288

; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

; SEQ ID NO 14338

; LENGTH: 175

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO AC021609.3

; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.48  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.58  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.52  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.57  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.79  
; OTHER INFORMATION: SWISSPROT HIT: O61608, EVALUOR 4.00e-04  
; OTHER INFORMATION: EST\_HUMAN HIT: AA085543.1, EVALUOR 7.00e-94  
; OTHER INFORMATION: NT HIT: g113325067, EVALUOR 5.00e-94  
US-10-029-386-14338

Query Match 8.3%; Score 175; DB 16; Length 175;  
Best Local Similarity 100.0%; Pred. No. 1.4e-84;

Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1770 AAAAGAGCTCAGACATTTCTTAAGCATGGATCTTAATCATCTAAAGTTTCCCTTC 1829  
DB 175 AAAAGAGCTCAGACATTTCTTAAGCATGGATCTTAATCATCTAAAGTTTCCCTTC 116  
QY 1830 AAGAGATGCTCTGTGGGAGAGAGAAAGCCCAAGAAAGTATGACAGCAACATCCA 1889  
DB 115 AAGAGATGCTCTGTGGGAGAGAGAAAGCCCAAGAAAGTATGACAGCAACATCCA 56  
QY 1890 GCTTCATGGCCAGACAGTGGGAGAAATCTCTCCAGAGAAAGGCCATATTAT 1944  
DB 55 GCTTCATGGCCAGACAGTGGGAGAAATCTCTCCAGAGAAAGGCCATATTAT 1

Search completed: August 27, 2005, 17:33:25

Job time : 903.401 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 23:18:31 ; Search time 4546.04 Seconds  
(without alignments)  
17558.328 Million cell updates/sec

Title: US-09-371-347a-43

Perfect score: 2097  
Sequence: 1 atgagagaggttcgtctact.....ttcagatatttgcataaa 2097

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-Processing: Listing first 45 summaries

Database : EST.\*

1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_hgc.\*  
4: gb\_est3.\*  
5: gb\_est4.\*  
6: gb\_est5.\*  
7: gb\_est6.\*  
8: gb\_gest.\*  
9: gb\_gest2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1592	75.9	3100	3	BC062577 Homo sapi
2	905	43.2	3143	3	BC035977 Homo sapi
3	719	34.3	908	5	BX348674 BX348674
4	638	30.7	874	4	BM801462 AGENCOURT
5	623	29.7	646	7	CN260357 170004241
6	586	27.9	852	5	BQ431497 AGENCOURT
7	565	26.9	865	1	AU279788 AU279788
8	543	25.9	877	1	AU124440 AU124440
9	531	25.3	1061	5	BQ218755 AGENCOURT
10	507	24.2	834	5	BU941078 AGENCOURT
11	470	22.4	521	6	CH164340 K-EST0225
12	466	22.2	826	4	BI772430 603055786
13	461	22.0	776	6	CB997527 AGENCOURT
14	455	21.7	822	1	AU132586 AU132586
15	448	21.4	591	2	AM965709 EST37782
16	446	21.3	818	6	CD559384 AGENCOURT
17	434	20.7	591	4	BI025283 RCS-MT025
18	406	19.4	710	5	BU570323 AGENCOURT
19	399	19.0	974	5	BX375211 BX375211
20	384	18.3	527	4	BI025277 RCS-MT025
21	374	17.8	579	7	CN260360 170006001
22	367	17.5	642	2	BF346446 602020302
23	361	17.2	692	7	CN260359 170004706
24	359	17.1	499	6	CD704108 EST20635

25	354	16.9	386	1	AA279726 ZB92410.1
26	351	16.7	839	2	BG531787 BG531787
27	340	16.2	526	4	AM952883 EST364953
28	337	16.1	818	7	CP995233 AGENCOURT
29	335	16.0	413	2	BF810368 RCS-C1041
30	335	16.0	413	2	BF810479 RCS-C1014
31	332	15.8	366	1	AA085543 ZN4411.1
32	331	14.8	478	4	BM754488 K-EST0031
33	309	14.7	685	4	BM049352 603626120
34	308	14.7	620	7	CK002453 AGENCOURT
35	302	14.4	440	4	BG877205 QV3-HT046
36	292	13.9	528	2	BE301292 B689B07.X
37	290	13.8	395	4	BM838530 K-EST0114
38	272	13.0	301	1	AL704780 DKFP0686M
39	264	12.6	366	6	CB298361 220019.1
40	257	12.3	366	2	BF808461 QV1-C1017
41	257	12.3	368	1	AA355001 EST63417
42	252	12.0	324	1	AA469901 ZL94D04.1
43	249	11.9	664	7	CR768694 DKFP459K
44	249	11.9	667	7	CR770923 DKFP469N
45	249	11.9	767	7	CR557482 DKFP469K

#### ALIGNMENTS

RESULT 1	BC062577	3100 bp	mRNA	linear	HTC 25-NOV-2003
LOCUS	Homo sapiens	cdna clone IMAGE:5189058, containing frame-shift errors.			
DEFINITION	BC062577.1	GI:38511756			
ACCESSION	BC062577				
VERSION	BC062577.1				
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 3100)				
AUTHORS	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buettow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heish, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stabile, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uedl, T.B., Toshnyk, S., Carninci, P., Prange, C., Kaba, S.S., Loggellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McGowan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hu, L., Hult, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywicki, M.I., Skalek, U., Smal, D.E., Scherch, A., Schein, U.E., Jones, S.J., and Marra, M.A.				
TITLE	human and mouse cDNA sequences				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	99 (26), 16899-16903			(2002)
MEDLINE	22388257				
PUBMED	12477932				
REFERENCE	2 (bases 1 to 3100)				
AUTHORS	Strausberg, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (24-NOV-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550, USA				
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>				
COMMENT	Contact: MGC help desk Email: <a href="mailto:cgabs-rcmail.nih.gov">cgabs-rcmail.nih.gov</a> Tissue Procurement: Life Technologies, Inc.				

cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: nisc\_ungcngrl.nih.gov  
Ahter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
Blakesley, R.W., Boufard, G.G., Breen, K., Brinkley, C., Brooks, S.,  
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,  
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Lurie, P., Legaspi, R.,  
Maduro, O.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,  
McDonnell, J., Pearson, R., Stantipop, S., Thomas, P.J., Touchman, J.W.,  
Tsougeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
Young, A., Zhang, L.-H., and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>  
Series: IRAC Plate: 135 Row: e Column: 21  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA 91: 4505278  
This clone has the following problem: frame shifted.

## FEATURES

## source

1. 3100  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5189058"  
/isssue type="Colon, Kidney, Stomach, adult, whole pooled"  
/clone\_id="NIH\_MGC\_116"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"

## ORIGIN

Query Match 75.9%; Score 1592; DB 3; Length 3100;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1812; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
283 GGTCTCGGTATTCAGAAATACACTTCTTTCATGAGGGGGAAGATATGTAACGA 342  
172 GGTCCTCGGTATTCAGAAATACACTTCTTTCATGAGGGGGAAGATATGTAACGA 231  
343 CTTCAGAGCTTGAAGCCCGCATTTTCTATGACCTGAGATGACATGATGATGATGAT 402  
232 CTTCAGAGCTTGAAGCCCGCATTTTCTATGACCTGAGATGACATGATGATGATGAT 291  
403 TTGAACTTGTGTTGACCGGTGATGCTGACCTGACCGCCCTCAGAAAGCATTTT 462  
292 TTGAACTTGTGTTGACCGGTGATGCTGACCTGACCGCCCTCAGAAAGCATTTT 351  
463 AGGTCAAGCAGAGCAGAGAGATAGTGGCGCACTCCCGGTGAGCATCACTGCATCC 522  
352 AGGTCAAGCAGAGCAGAGAGATAGTGGCGCACTCCCGGTGAGCATCACTGCATCC 411  
523 TTGAGCAGACCTTGTGAGTCACTGACCTGACATTTGAATTCGAGTTCGAGTTCTG 582  
412 TCGAGCAGACCTTGTGAGTCACTGACCTGACATTTGAATTCGAGTTCGAGTTCTG 471  
583 AGATTCATGATTCAGAGAGAGAGATTCGAGGTTTGAAGCAAAATGACGTGACAGC 642  
472 AGATTCATGATTCAGAGAGAGAGATTCGAGGTTTGAAGCAAAATGACGTGACAGC 531  
643 AACCAATCAATGTTGATTAATTAAGATTCGAGTTCCTCACTTACCGGTTCGATCCCA 702  
532 AACCAATCAATGTTGATTAATTAAGATTCGAGTTCCTCACTTACCGGTTCGATCCCA 591  
703 CTCTCAAGGCTCTCGAATATTCCTGTTTACCCCAAGATTAATTAAGGATTCAGTTCG 762  
592 CTCTCAAGGCTCTCGAATATTCCTGTTTACCCCAAGATTAATTAAGGATTCAGTTCG 651  
763 CAGGAGCTCTTGGCAGAGAGAGAGATTCGAGTTCGATTCAGGATTCAGGATTCAGTTCG 822  
652 CAGGAGCTCTTGGCAGAGAGAGAGATTCGAGTTCGATTCAGGATTCAGGATTCAGTTCG 711

QY 823 CAAATGCCAATTTCAAAAGGAGTTCACTTACTAGCAATGATGCCATTAACCACTCTG 882  
DB 712 CAAATGCCAATTTCAAAAGGAGTTCACTTACTAGCAATGATGCCATTAACCACTCTG 771  
QY 883 CTGATGAAATTTGAGCAATTTCAAAATGAGATTTTCTATCAAGCTTGAAGTCTTCAAG 942  
DB 772 CTGATGAAATTTGAGCAATTTCAAAATGAGATTTTCTATCAAGCTTGAAGTCTTCAAG 831  
QY 943 GTGATCTGCTTACAGTATTTCTGAGATTAACAAGCTTCTCAAGAGCTGACCTTGA 1002  
DB 832 GTGATCTGCTTACAGTATTTCTGAGATTAACAAGCTTCTCAAGAGCTGACCTTGA 891  
QY 1003 GATTAAGAGAGAGCTGCTCTTTTGAATTAAGGAGAGAGAGAGAGAGAGAGCT 1062  
DB 892 GATTAAGAGAGAGCTGCTCTTTTGAATTAAGGAGAGAGAGAGAGAGAGAGAGCT 950  
QY 1063 ACCTTTCCCAAGATTAATCTGAGGATTTCTCTCAAGTCAATTTTCAAGTCTTCTG 1122  
DB 951 ACCTTTCCCAAGATTAATCTGAGGATTTCTCTCAAGTCAATTTTCAAGTCTTCTG 1010  
QY 1123 GAAATCCGAGCAATTTCTTAATAAGATTTTGGAGAGCTTGTGACTATACAGTAC 1182  
DB 1011 GAAATCCGAGCAATTTCTTAATAAGATTTTGGAGAGCTTGTGACTATACAGTAC 1070  
QY 1183 AGTCTGAAAG 1242  
DB 1071 AGTCTGAAAG 1130  
QY 1243 CGCTTTGTAAG 1302  
DB 1131 CGCTTTGTAAG 1190  
QY 1303 CAGCCAGCACTAGCTCTCTGCTGAGCAATCTTCTTAACCTTCAACCCAGAGAGATTCG 1362  
DB 1191 CAGCCAGCACTAGCTCTCTGCTGAGCAATCTTCTTAACCTTCAACCCAGAGAGATTCG 1250  
QY 1363 TGTGCAAGCTCAAGTTATTTTCAACCCAGAGAGAGAGAGAGAGAGAGAGAGAG 1422  
DB 1251 TGTGCAAGCTCAAGTTATTTTCAACCCAGAGAGAGAGAGAGAGAGAGAGAGAG 1310  
QY 1423 TTTCTGCTACAG 1482  
DB 1311 TTTCTGCTACAG 1370  
QY 1483 TTGTTGTTGCTTCAAGTCTTCAAGCAATCATGATTCATGATTCATGATTCATGAT 1542  
DB 1371 TTGTTGTTGCTTCAAGTCTTCAAGCAATCATGATTCATGATTCATGATTCATGAT 1430  
QY 1543 GCTCTGCTCTTAAGATATCATCTCTCTGAGCAACCAATCTTCTCAAGTTCATGAT 1602  
DB 1431 GCTCTGCTCTTAAGATATCATCTCTCTGAGCAACCAATCTTCTCAAGTTCATGAT 1490  
QY 1603 GACCCCTCAATCCCAATTAATGATGATGATGATGATGATGATGATGATGATGAT 1662  
DB 1491 GACCCCTCAATCCCAATTAATGATGATGATGATGATGATGATGATGATGATGAT 1550  
QY 1663 TTCTTCAACATTAAG 1722  
DB 1551 TTCTTCAACATTAAG 1610  
QY 1723 TGGTTGTTTGGCTGAG 1782  
DB 1611 TGGTTGTTTGGCTGAG 1670  
QY 1783 CATTCCTTAAG 1842  
DB 1671 CATTCCTTAAG 1730  
QY 1843 GTTGGAG 1902  
DB 1731 GTTGGAG 1790

QY 1903 CAGTGGCCAGATCTCTCCAGAGAAAGCCATATTATGATGAGATGCAAG 1962  
DB 1791 CAGGTGGGAGATCTCTCCAGAGAAAGCCATATTATGATGAGATGCAAG 1850  
QY 1963 AATATGGCCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2022  
DB 1851 AATATGGCCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1910  
QY 2023 GAAAACTGAGCATGATGATGATGATGATGATGATGATGATGATGATGAT 2082  
DB 1911 GAAAACTGAGCATGATGATGATGATGATGATGATGATGATGATGATGAT 1970  
QY 2083 GATATTGGTCATTA 2097  
DB 1971 GATATTGGTCATTA 1985

RESULT 2  
BC035977 3143 bp mRNA linear HTC 20-SEP-2002  
LOCUS BC035977 Home sapiens, clone IMAGE:461253, mRNA.  
ACCESSION BC035977  
VERSION BC035977.1 GI:23243305  
KEYWORDS HTC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 3143)  
Strausberg, R.  
Direct Submission  
Submitted (31-JUL-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
Contact: MGC help desk  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: CLONTECH  
CDNA Library Preparation: CLONTECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www-ehgc.stanford.edu>  
Contact: (Dickson, Mark) [med@paxli.stanford.edu](mailto:med@paxli.stanford.edu)  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
R. M.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
Series: IRML Plate: 41 Row: g Column: 2  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 4505278  
This clone has the following problem: frame shifted.  
Location/Qualifiers  
1..3143  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:461253"  
/tissue\_type="Kidney"  
/clone\_lib="NIH MGC\_75"  
/lab\_host="DH10B"  
/note="Vector: pDNR-LIB"

ORIGIN  
Query Match 43.2%; Score 905; DB 3; Length 3143;  
Best Local Similarity 99.7%; Pred. No 0;  
Matches 1055; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGAGAGAGTTTCTGTACTATATGCTACACAGCAGGAGCAAGCCATCGCAGAA 60  
DB 52 ATGAGAGAGTTTCTGTACTATATGCTACACAGCAGGAGCAAGCCATCGCAGAA 111

QY 61 GAAATGTGAGCAAGCTGTGTATGATGATGATGATGATGATGATGATGATGATGAT 120  
DB 112 GAAATGTGAGCAAGCTGTGTATGATGATGATGATGATGATGATGATGATGATGAT 171  
QY 121 TCCGATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
DB 172 TCCGATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 231  
QY 181 GGCACCGGAGACCCAGCCAGCAGCCGAGGATGATGATGATGATGATGATGATGAT 240  
DB 232 GGCACCGGAGACCCAGCCAGCAGCCGAGGATGATGATGATGATGATGATGATGAT 291  
QY 241 CTGCGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
DB 292 CTGCGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 351  
QY 301 TACACCTACTTTTGCATGAGGAGGAGGATGATGATGATGATGATGATGATGATGAT 360  
DB 352 TACACCTACTTTTGCATGAGGAGGAGGATGATGATGATGATGATGATGATGATGAT 411  
QY 361 CGGCAATTTTATGACACTGATGATGATGATGATGATGATGATGATGATGATGAT 420  
DB 412 CGGCAATTTTATGACACTGATGATGATGATGATGATGATGATGATGATGATGAT 471  
QY 421 CCGTGATGTTGCTGACCTGCGCAGCCCTCAGAAAGCATTTTAGGTCAAGAGACAA 480  
DB 472 CCGTGATGTTGCTGACCTGCGCAGCCCTCAGAAAGCATTTTAGGTCAAGAGACAA 531  
QY 481 GAGGAGATTAAGTGGGCGCATCTCCGGTGATGATGATGATGATGATGATGATGAT 540  
DB 532 GAGGAGATTAAGTGGGCGCATCTCCGGTGATGATGATGATGATGATGATGATGAT 591  
QY 541 AAGTCAGACCTGCTACACTGATGATGATGATGATGATGATGATGATGATGATGAT 600  
DB 592 AAGTCAGACCTGCTACACTGATGATGATGATGATGATGATGATGATGATGATGAT 651  
QY 601 AGAAGAGATTTGAGGTTTGAAGCAAAATGAGTGAACAGCAACCAATCCAAATGTTGA 660  
DB 652 AGAAGAGATTTGAGGTTTGAAGCAAAATGAGTGAACAGCAACCAATCCAAATGTTGA 711  
QY 661 ATTGAAGACTTTGAGTCTCACTTACCCGCTGATGATGATGATGATGATGATGATGAT 720  
DB 712 ATTGAAGACTTTGAGTCTCACTTACCCGCTGATGATGATGATGATGATGATGATGAT 771  
QY 721 AATATTCTGTTTACCCCAAGATTTTACAGATGATGATGATGATGATGATGATGAT 780  
DB 772 AATATTCTGTTTACCCCAAGATTTTACAGATGATGATGATGATGATGATGATGAT 831  
QY 781 GAGGAAGCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
DB 832 GAGGAAGCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 891  
QY 841 GCGATTCAACTTACTACGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
DB 892 GCGATTCAACTTACTACGATGATGATGATGATGATGATGATGATGATGATGATGAT 951  
QY 901 TCAATATCAAGCTTTTCTATGAGCTTGAAGATGATGATGATGATGATGATGATGAT 960  
DB 952 TCAATATCAAGCTTTTCTATGAGCTTGAAGATGATGATGATGATGATGATGATGAT 1011  
QY 961 GATTCTGAGGTACAAAGCTTCCAAAGACGCTGAGCTTGAAGATTAAGAGAGACCTGC 1020  
DB 1012 GATTCTGAGGTACAAAGCTTCCAAAGACGCTGAGCTTGAAGATTAAGAGAGACCTGC 1071  
QY 1021 GTCTTTTGAATAAAGCAGACACAAAGAAAGG 1058  
DB 1072 GTCTTTTGAATAAAGCAGACACAAAGAAAGG 1109

RESULT 3  
BX348674 908 bp mRNA linear EST 08-APR-2004  
LOCUS BX348674

DEFINITION BX348674 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
CDNA clone CS0DC010Y11 5-PRIME, mRNA sequence.  
ACCESSION BX348674  
VERSION BX348674.1 GI:30375301  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 908)  
AUTHORS Li, W.-B., Gruber, C., Jessee, J., and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Creteil, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 3392.f  
For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?cs=CSDBAG006ZB02\_CS00490\_1&c=3392.f

FEATURES  
source  
Location/Qualifiers  
1..908  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DC010Y11"  
/issue\_type="NEUROBLASTOMA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN  
Query Match 34.3%; Score 719; DB 5; Length 908;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 719; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 674 AGTCTCACTTACCGGTTGGTACCCCACTCTCAAGCCTCTGTAATTCCTGTT 733  
DB 28 AGTCTCACTTACCGGTTGGTACCCCACTCTCAAGCCTCTGTAATTCCTGTT 87  
QY 734 TACCCCGAATATTTACAGTATCTGCGAGGCTCTTGGCCAGAGAAAGCCAG 793  
DB 88 TACCCCGAATATTTACAGTATCTGCGAGGCTCTTGGCCAGAGAAAGCCAG 147  
QY 794 TATCTGACTTACAGAGATCCAGTTTTCAGAGCCAAATTTCAAGGAGTTCACCTTA 853  
DB 148 TATCTGACTTACAGAGATCCAGTTTTCAGAGCCAAATTTCAAGGAGTTCACCTTA 207  
QY 854 CTACGAATGATCCATATAAAACACTCTGCTGATGATGATGATGATGATGATGATGAT 913  
DB 208 CTACGAATGATCCATATAAAACACTCTGCTGATGATGATGATGATGATGATGATGAT 267  
QY 914 TTTTCTATCAGCTGAGAGATGCTTCAAGCTGATGCTTCAAGCTGATGATGATGATGATGAT 973  
DB 268 TTTTCTATCAGCTGAGAGATGCTTCAAGCTGATGCTTCAAGCTGATGATGATGATGATGAT 327  
QY 974 AAAGCTACTCCAAAGACTGAGCTGTAAGATTAAGAGAGACATGCGTCTTTGAAA 1093  
DB 328 AAAGCTACTCCAAAGACTGAGCTGTAAGATTAAGAGAGACATGCGTCTTTGAAA 387  
QY 1034 TAAAGGAGACAAAG 1093  
DB 388 TAAAGGAGACAAAG 447  
QY 1094 CTCCTCAGTTATTTTACCTGCTGCTTGAATCCAGAGCAATTCCTTAAAAAGGCAATTTT 1153

DB 448 CTCCTCAGTTATTTTACCTGCTGCTTGAATCCAGAGCAATTCCTTAAAAAGGCAATTTT 507  
QY 1154 TGGAGCCCTTGTGAGTATACAGTACAGAGTGTGAAAGCGCAGGCTTACAGAGAGCTGT 1213  
DB 508 TGGAGCCCTTGTGAGTATACAGTACAGAGTGTGAAAGCGCAGGCTTACAGAGAGCTGT 567  
QY 1214 GCAGTAAACAGAGGAGCGAGCCGATTTATGCGGCTTTGTAGAGATGCTGCTGCTTGT 1273  
DB 568 GCAGTAAACAGAGGAGCGAGCCGATTTATGCGGCTTTGTAGAGATGCTGCTGCTTGT 627  
QY 1274 TGGATCTCCCTCGCTTCCCTTCCCTTGGCAGCACCATCTGCTGCTGCAATC 1333  
DB 628 TGGATCTCCCTCGCTTCCCTTCCCTTGGCAGCACCATCTGCTGCTGCAATC 687  
QY 1334 TTCTTAACCTTCAACCCAGACCATATTCGTGCAAGCTCAAGTTATTTACCCAGGA 1392  
DB 688 TTCTTAACCTTCAACCCAGACCATATTCGTGCAAGCTCAAGTTATTTACCCAGGA 746

RESULT 4  
BM801462 874 bp mRNA linear EST 05-MAR-2002  
LOCUS  
DEFINITION AGENCOURT\_6459212 NIH\_MGC\_88 Homo sapiens cDNA clone IMAGE:5560477  
5', mRNA sequence.  
ACCESSION BM801462  
VERSION BM801462.1 GI:19118285  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 874)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapds-r@mail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LHAM12286 row: 1 column: 14  
High quality sequence stop: 710.

FEATURES  
source  
Location/Qualifiers  
1..874  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5560477"  
/issue\_type="duodenal adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_88"  
/note="Organ: small intestine; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.767 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC Library."

ORIGIN  
Query Match 30.4%; Score 638; DB 4; Length 874;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 738; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGAGAGTTTCTGTTACTATATGCTACACAGAGGAGACGCAAGCCATGCGAGAA 60  
DB 50 ATGAGAGAGTTTCTGTTACTATATGCTACACAGAGGAGACGCAAGCCATGCGAGAA 109  
QY 61 GAAATGTGAGAGAGCTGTGTGATGATGATTTTCTGAGATCTTCACTATATTAGGAA 120







Db 241 AGCTACCTTACCCGACATATACCTGCGGGAGTGTCTCTCCAGTTTCAATTTTTCAGTGTG 300

QY 1119 TCTTGAATCCGAGCAATTCCTAAAAGGCAATTTTGGAGCCCTTGGAGCTATACGAG 1178

Db 301 TCTTGAATCCGAGCAATTCCTAAAAGGCAATTTTGGAGCCCTTGGAGCTATACGAG 360

QY 1179 TGACAGTGTGTAAGGCGAGGCTACAGAGGCTGTGTCAGTAAACAAGGGGCGAGCCGATTA 1238

Db 361 TGACAGTGTGTAAGGCGAGGCTACAGAGGCTGTGTCAGTAAACAAGGGGCGAGCCGATTA 420

QY 1239 TAGCGGCTTGTATACGAGATGCTGCTGCTGCTTGTGATCTCTCTGCTTCCCTTC 1298

Db 421 TAGCGGCTTGTATACGAGATGCTGCTGCTGCTTGTGATCTCTCTGCTTCCCTTC 480

QY 1299 TTGCGAGGCAACACCTACGCTCTGCTGCAACCTCTCTAAACCTTCAACCGACGACATTA 1358

Db 481 TTGCGAGGCAACACCTACGCTCTGCTGCAACCTCTCTAAACCTTCAACCGACGACATTA 540

QY 1359 TTGCGTGTGCAAGCTCAAGTTTATTT 1383

Db 541 TTGCGTGTGCAAGCTCAAGTTTATTT 565

RESULT 8

LOCUS AU124440 877 bp mRNA linear EST 01-AUG-2002

DEFINITION AU124440 NT2RM4 Homo sapiens cDNA clone NT2RM400010 5', mRNA

ACCESSION AU124440

VERSION AU124440

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS Ota, T., Wakamatsu, A., Ozawa, M., Ishii, S., Saito, K., Yamamoto, J., Nakamura, Y., Nishikawa, T., Nagai, T., Suzuki, Y., Sugano, S. and Isegai, T.

TITLE HRI human cDNA project (Ota, T., Wakamatsu, A., Ozawa, M., Ishii, S., Saito, K., Yamamoto, J., Nakamura, Y., Nishikawa, T., Nagai, T., Suzuki, Y., Sugano, S., Isegai, T.)

JOURNAL Unpublished (2000)

COMMENT Contact: Takao Isegai

Genomics Laboratory

Helix Research Institute

1532-3 Yana, Kisarazu, Chiba 292-0812, Japan

Tel: 81-438-52-3975

Fax: 81-438-52-3986

Email: genomics@hri.co.jp

HRI human cDNA project; 5', & 3'-end one pass sequencing; Helix Research Institute; cDNA library construction; Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

FEATURES

source

1.877

Location/Qualifiers

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="NT2RM400010"

/cell\_type="teratocarcinoma"

/cell\_line="NT2"

/clone\_1fb="NT2RM4"

/note="Vector: pME18SFL3; mRNA from uninduced NT2 neuronal precursor cells"

ORIGIN

Query Match 25.9%; Score 543; DB 1; length 877;

Best Local Similarity 100.0%; Pred.No.12e-286;

Matches 543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1143 AAGGCAATTTTTCGAGCCCTTGTGACTATACAGTGAAGTGAAGGCGAGGCT 1202

Db 192 AAGGCAATTTTTCGAGCCCTTGTGACTATACAGTGAAGTGTGTAAGGCGAGGCT 251

QY 1203 ACAGAGCTGTGACAGTAAACAAGGGGCGAGCCGATTAATAGCCGCTTGTACAGATGCTG 1262

Db 252 ACAGAGCTGTGACAGTAAACAAGGGGCGAGCCGATTAATAGCCGCTTGTACAGATGCTG 311

QY 1263 TGCCTGTGTGTGATATCTCTCTCTGCTTCCCTTCTTCCAGCCACAACCTGCTCTCT 1322

Db 312 TGCCTGTGTGTGATATCTCTCTCTGCTTCCCTTCTTCCAGCCACAACCTGCTCTCT 371

QY 1323 GCTGCAACATCTTCTCTAAACCTTCAACCCAGACCATATGCTGTGCAAGCTCAAGTTAT 1382

Db 372 GCTGCAACATCTTCTCTAAACCTTCAACCCAGACCATATGCTGTGCAAGCTCAAGTTAT 431

QY 1383 TCACCGAAGAAAGCTCCATTTGTCTTCAACATGTGGAATTTCTGTACTGACCAAC 1442

Db 432 TCACCGAAGAAAGCTCCATTTGTCTTCAACATGTGGAATTTCTGTACTGACCAAC 491

QY 1443 AAGAGTTCTGCGGAAGGAGATGTATACAGGCTGAGCTGTGTTGTTCAAGTTCT 1502

Db 492 AAGAGTTCTGCGGAAGGAGATGTATACAGGCTGAGCTGTGTTGTTCAAGTTCT 551

QY 1503 TCAGCCAAACATATCATGCTATCCATGAGACAGCGGAAAGCCCTGCTCTTAAGATATC 1562

Db 552 TCAGCCAAACATATCATGCTATCCATGAGACAGCGGAAAGCCCTGCTCTTAAGATATC 611

QY 1563 CATCTCTCTGCAACAAATTTCTTCCACTTACAGATGACCCCTCAATCCCATCAT 1622

Db 612 CATCTCTCTGCAACAAATTTCTTCCACTTACAGATGACCCCTCAATCCCATCAT 671

QY 1623 AATGTGGTGTCCAGAACCGGATAGCCCGTTTATTTGGTTCTTACACATAGAGAGA 1682

Db 672 AATGTGGTGTCCAGAACCGGATAGCCCGTTTATTTGGTTCTTACACATAGAGAGA 731

QY 1683 ACT 1685

Db 732 ACT 734

RESULT 9

LOCUS BQ218755 1061 bp mRNA linear EST 02-MAY-2002

DEFINITION AGENCOURT 7565843 NIH\_MGC\_92 Homo sapiens cDNA clone IMAGE:6041670 5', mRNA sequence.

ACCESSION BQ218755

VERSION BQ218755

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS NIH-MGC <http://mhc.mci.nih.gov/>.

TITLE NIH-MGC

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>

Place: LAM13279 row: n column: 07

High quality sequence stop: 518.

Location/Qualifiers

1.1061

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6041670"

/tissue\_type="embryonal carcinoma, cell line"

/lab host="DH10B (phage-resistant)"  
/clone lib="NIH\_MGC\_92"  
/note="Organ: testis; Vector: PCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 2.5 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 25.3%; Score 531; DB 5; Length 1061;  
Best Local Similarity 100.0%; Pred. No. 5.2e-280;  
Matches 531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 995 AGCTTGAAATGAAAGAGACACTGCGTCTTTGAAATGAAAGGACAGACAAAGAGA 1054  
Db 1 AGCTTGAAATGAAAGAGACACTGCGTCTTTGAAATGAAAGGACAGACAAAGAGA 60

QY 1055 AAGGAGCTACCTTACCCAGCATTTACCTGCGGAGATTTCTTCCAGTTATTTTAACT 1114  
Db 61 AAGGAGCTACCTTACCCAGCATTTACCTGCGGAGATTTCTTCCAGTTATTTTAACT 120

QY 1115 GGTGCTTGAATCCGAGCAATTCCTTAAAGGCAATTTTTCGAGCCCTTGTGACTATA 1174  
Db 121 GGTGCTTGAATCCGAGCAATTCCTTAAAGGCAATTTTTCGAGCCCTTGTGACTATA 180

QY 1175 CCAAGTGAAGTGTGAAAGGAGGCTAGAGAGCTGTGCAATTAACAAGGGGAGCCG 1234  
Db 181 CCAAGTGAAGTGTGAAAGGAGGCTAGAGAGCTGTGCAATTAACAAGGGGAGCCG 240

QY 1235 ATTAATACCGCTTTGTAAGAGATGCTGTGCTGCTGTTGATCTCTCTGCTTCC 1294  
Db 241 ATTAATACCGCTTTGTAAGAGATGCTGTGCTGCTGTTGATCTCTCTGCTTCC 300

QY 1295 CTCTGTCGAGCCAGCACTGCTGCTGCTGCAATCTTCTTAACTTCAACCCAGAC 1354  
Db 301 CTCTGTCGAGCCAGCACTGCTGCTGCTGCAATCTTCTTAACTTCAACCCAGAC 360

QY 1355 CATATTCGTGTGCAAGTCAAGTTATTTCAACCCAGAAAGCTCATTTGTCTTACA 1414  
Db 361 CATATTCGTGTGCAAGTCAAGTTATTTCAACCCAGAAAGCTCATTTGTCTTACA 420

QY 1415 TTGTGGAATTTCTGTCTTACGCAACAAGAGTTTTCGGAAGAGATATGACAGCT 1474  
Db 421 TTGTGGAATTTCTGTCTTACGCAACAAGAGTTTTCGGAAGAGATATGACAGCT 480

QY 1475 GGCTGGCGCTGTGTTGCTTCACTTCTTCAAGCCAAATATGATGATCC 1525  
Db 481 GGCTGGCGCTGTGTTGCTTCACTTCTTCAAGCCAAATATGATGATCC 531

RESULT 10  
BU941078 834 bp mRNA linear EST 18-OCT-2002  
LOCUS  
DEFINITION AGNCOURT\_10540067 NIH\_MGC\_128 Homo sapiens cDNA clone  
IMAGE:6712893 5', mRNA sequence.  
ACCESSION BU941078  
VERSION BU941078.1 GI:24129897  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 834)  
NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: csapbs-remail.nih.gov  
Tissue Procurement: NCI  
cDNA Library Preparation: Michael Brownstein Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LMC3022 row: e column: 21  
High quality sequence stop: 586.  
Location/Qualifiers

## FEATURES

source

1. 834  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6712893"  
/tissue\_type="mixed (pool of 40 RNAs)"  
/lab\_host="DH10B (T1-phage-resistant)"  
/clone\_lib="NIH\_MGC\_128"  
/note="Vector: pDNR-LIB; Site 1: SfiI (ggccattagcgc);  
Site 2: SfiI (ggccgcctgcgc); Double-stranded cDNA was  
prepared from a pool of 40 cell line polyA+ RNAs (bladder  
- 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -  
4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,  
kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell -  
5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%,  
salivary gland - 1.3%, and skin - 2.3%). 5' and 3'  
adaptors were used in cloning as follows:  
5'-AAGAGTGTATCAAGCAGAGGCGCATTAAGCGCGG-3' and  
5'-ATTCTAGAGCGGAGGCGGCACATG-AT(30)NN-3'. Full-length  
enriched library was constructed using the Clontech  
Creator SMART Kit and size-selected to contain the >2 kb  
size fraction (other fractions present in NIH\_MGC\_126 and  
NIH\_MGC\_127). Library created in the laboratory of T.  
Uedlin, M.D., Ph.D. (NIH, NIH). Note: this is a NIH\_MGC  
Library."

## ORIGIN

Query Match 24.2%; Score 507; DB 5; Length 834;  
Best Local Similarity 99.7%; Pred. No. 8.3e-267;  
Matches 607; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 379 GGACATGCAAGATCTGTGATGTTTGAACCTTGTGAGCCGCGATGCTGAGCTC 438  
Db 3 GGACATGCAAGATCTGTGATGTTTGAACCTTGTGAGCCGCGATGCTGAGCTC 62

QY 439 TGGCCAGCCCTCGAAGAACTTTAGTCAAGCAGAGACAAGAGAGATTAATGCGCA 498  
Db 63 TGGCCAGCCCTCGAAGAACTTTAGTCAAGCAGAGACAAGAGAGATTAATGCGCA 122

QY 499 CTCGCGGTGCAATCACTGATCTTGAAGACAGACCTTGTGAAGTCAAGCTGTACAC 558  
Db 123 CTCGCGGTGCAATCACTGATCTTGAAGACAGACCTTGTGAAGTCAAGCTGTACAC 182

QY 559 ATTGAATCTCAAGTCAAGCTTCTGAGATTCGATGATTCAGAAAGAAATTTCTAGCTT 618  
Db 183 ATTGAATCTCAAGTCAAGCTTCTGAGATTCGATGATTCAGAAAGAAATTTCTAGCTT 242

QY 619 TTGAAGCAAAATGACAGTGAACAGCAACCAATCAATGTTGTAATTTGAAGCTTGAATCC 678  
Db 243 TTGAAGCAAAATGACAGTGAACAGCAACCAATCAATGTTGTAATTTGAAGCTTGAATCC 302

QY 679 TCACTTACCGGTTGCTGATCCCACTCTCAAGAGCTTCTGTAATATTCCTGTTTACC 738  
Db 303 TCACTTACCGGTTGCTGATCCCACTCTCAAGAGCTTCTGTAATATTCCTGTTTACC 362

QY 739 CCAAGATATTTAAGTACATCTGCAAGAGTCTTGGCCAGAGAGAAAGCCAAATATCT 798  
Db 363 CCAAGATATTTAAGTACATCTGCAAGAGTCTTGGCCAGAGAGAAAGCCAAATATCT 422

QY 799 GTGACTTACGACATCAAGTTTTCAGAGTCAAGTTTCAAGGCAAGTCACTTACTACG 858  
Db 423 GTGACTTACGACATCAAGTTTTCAGAGTCAAGTTTCAAGGCAAGTCACTTACTACG 482

QY 859 AATGATCCATTAACCAACTCTGCTGTGTAATTTGACATTTCAATACAGACTTTTCC 918  
Db 483 AATGATCCATTAACCAACTCTGCTGTGTAATTTGACATTTCAATACAGACTTTTCC 542

QY 919 TATACGCTTGAAGATGCTTCAAGCTGATCTGCCCTTACAGTATTTCTGAGTACAAAGC 978

Db	543	TATAGCTGAGATGCTTGAAGGTATCTGCCTTAACAGTATTTGAGGTACAAGC	602
Qy	979	CTACTCCA 987	
Db	603	CTACTCCA 611	
RESULT 11			
LOCUS	CB164340		
DEFINITION	K-EST0225498 L17N670205n1 Homo sapiens CDNA clone		
VERSION	L17N670205n1-39-F02 5', mRNA sequence.		
KEYWORDS	CB164340		
SOURCE	CB164340.1 GI:28150466		
ORGANISM	EST.		
REFERENCE	Homo sapiens (human)		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
TITLE	1 (bases 1 to 521)		
JOURNAL	Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.B., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.		
COMMENT	21C Frontier Korean EST Project 2001		
FEATURES	Unpublished (2002)		
Source	Contact: Kim YS		
	Genome Research Center		
	52 Eosun-dong Yuseong-gu, Daejeon 305-333, South Korea		
	Tel: +82-42-860-4470		
	Fax: +82-42-860-4409		
	Email: yongsung@mail.krdb.re.kr		
	Plate: 39 row: F column: 02		
	High quality sequence stop: 521.		
	location/Qualifiers		
	1..521		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="L17N670205n1-39-F02"		
	/sex="F"		
	/lab_host="Top10F"		
	/clone_idb="L17N670205n1"		
	/note="Organ: Liver; Vector: pTY71-Pac; Site 1: EcoRI; Site 2: NotI; The library was contributed by the Soares laboratory and it was constructed as described by Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome Research 6(9): 791-806. RNA was prepared from harvested cell culture."		
ORIGIN			
Query Match	22.4%	Score 470; DB 6; Length 521;	
Best Local Similarity	99.8%;	Fred. No. 1.9e-246;	
Matches	520; Conservative	0; Mismatches 1; Indels 0; Gaps 0;	
Qy	1540	AAAGCCGGGCTCCGATATATCCATCTCTCTGGAACAAGAAATTTTCCACTTACCA	15999
Db	1	AAAGCCCTGGCTCTTAAGATATCCATCTCTCTGGAACAAGAAATTTTCCACTTACCA	60
Qy	1600	GATGACCCCTCAATCCCATCATATATATGTTGGGTCGAGAACCGGATAGCCCGTTTAT	16598
Db	61	GATGACCCCTCAATCCCATCATATATATGTTGGGTCGAGAACCGGATAGCCCGTTTAT	120
Qy	1660	GGGTTTCCATCAACATAGAGAGAAACTCCAAAGAACCAACCCGATGGAATTTTGGAGCA	17198
Db	121	GGGTTTCCATCAACATAGAGAGAAACTCCAAAGAACCAACCCGATGGAATTTTGGAGCA	180
Qy	1720	ATGTGGTTGTTTGGTGGCTGAGGATTAAGATAGGATTAATCTATTTCAGAAAAGAGCTC	17797
Db	181	ATGTGGTTGTTTGGTGGCTGAGGATTAAGATAGGATTAATCTATTTCAGAAAAGAGCTC	240
Qy	1780	AGACATTTCTTAAAGCATGGATCTTAACTCATCTAAAGGTTTTCTTCTTCAGAGAGTCT	18399

Db		241	AGATATTTCCTTAAGCATGGGATCTTTAACTCACTTAAAGGTTTCCTCTCAAGAGATGCT	300
OY		1840	CCTGTGGGAGGAGGAAGCCCGACGAAAGTAGTATACAAGACAACTCCAGCTTCATAGGC	1899
Db		301	CCTGTGGGAGGAGGAAGCCCGACGAAAGTAGTATACAAGCAACTCCAGCTTCATAGGC	360
OY		1900	CAGCAGGTGGCGAAGAACCTCCTCCACAGAGAAGGCCCATTTATGTGTGTGAGATGCA	1959
Db		361	CAGCAGGTGGCGAAGAACCTCCTCCACAGAGAAGGCCCATTTATGTGTGTGAGATGCA	420
OY		1960	AAGAATATGGCCAAGATGTATCATGATAGCCCTTGTCMAATATATAGCAAAGAGTTGGA	2019
Db		421	AAGAATATGGCCAAGATGTATCATGATAGCCCTTGTCMAATATATAGCAAAGAGTTGGA	480
OY		2020	GTTGAAAAAAGCTAGAGCAATGAAAACCCTGGCCACTTTAAA	2060
Db		481	GTTGAAAAAAGCTAGAGCAATGAAAACCCTGGCCACTTTAAA	521
RESULT 12				
Bt772430		826 bp	mRNA	linear EST 25-SEP-2001
LOCUS		603055786F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5205285 5'		
DEFINITION		mRNA sequence.		
ACCESSION		B1772430		
VERSION		B1772430.1 GI:15764008		
KEYWORDS		EST.		
SOURCE		Homo sapiens (human)		
ORGANISM		Homo sapiens		
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE		NIH-MGC http://mgs.nci.nih.gov/.		
JOURNAL		National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT		Unpublished (1999) Contact: Robert Strauberg, Ph.D. Email: cgabbs@emall.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov Plate: LHAM1514 row: 1 column: 22 High quality sequence stop: 824.		
FEATURES		Location/Qualifiers		
Source		1..826		
		/organism="Homo sapiens"		
		/mol_type="mRNA"		
		/db_xref="taxon:9606"		
		/clone="IMAGE:5205285"		
		/lab_host="DH10B"		
		/clone_id="NIH_MGC_122"		
		/note="Organ: pooled lung and spleen; Vector: pCMV-Sport6 Site 1: NotI; Site 2: BcoRI (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens." Library is oligo-dT primed and directionally cloned (BcoRI site is destroyed upon cloning). Average insert size 1.4 kb. Insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH_MGC library."		
ORIGIN				
Query Match		22.2%	Score 466; DB 4; Length 826;	
Best Local Similarity		99.5%	Pred. No. 3.2e-244;	
Matches 616; Conservative		0; Mismatches 3; Indels 0; Gaps 0;		
OY		1	ATGAGAGGTTTCTTACTATATGCTACACAGCAGGACAGCGAAGGCCATGTGGCAAA	60

```
Db 53 ATGAGAGGTTTCTGTTACTATATGCTACACAGACGGGACAGCGAAAGCCATCGACAA 112
Qy 61 GAAATGTGAGCAAGCTGTGTGATACATGATATTTCTGCAGATCTTCACTATATTAGTAA 120
Db 113 GAAATATGTAGCAAGCTGTGTGATACATGATATTTCTGCAGATCTTCACTATATTAGTAA 172
Qy 121 TCCGATTAAGTATGACCTTAACCAACCGAAACGCTCTCTTGTGTGTGTTTCTACACG 180
Db 173 TCCGATTAAGTATGACCTTAACCAACCGAAACGCTCTCTTGTGTGTGTTTCTACACG 232
Qy 181 GGCACCGGAGACCCACCCGACACAGCCCGAAGTTGTTAAGGAAATACAGAACCAACA 240
Db 233 GGCACCGGAGACCCACCCGACACAGCCCGAAGTTGTTAAGGAAATACAGAACCAACA 292
Qy 241 CTGCGCGTGTATTTCTTGTCTCACTGCGGTATGAGTTAGCTGCGTCTCGGTGATTGAA 300
Db 293 CTGCGCGTGTATTTCTTGTCTCACTGCGGTATGAGTTAGCTGCGTCTCGGTGATTGAA 352
Qy 301 TACACCTACTTTTGCATTTGGGGGAGATTAATGATTAACGACTTCAAGCTTGGAGCC 360
Db 353 TACACCTACTTTTGCATTTGGGGGAGATTAATGATTAACGACTTCAAGCTTGGAGCC 412
Qy 361 CGGATTTCTATGACATGACATGACATGACATGACATGACATGACATGACATGACATGAC 420
Db 413 CGGATTTCTATGACATGACATGACATGACATGACATGACATGACATGACATGACATGAC 472
Qy 421 CCGTGATTTGCTGACCTCTGCGCACGCTCTCAAGAAAGATTTAGTCAACAGACAGACAA 480
Db 473 CCGTGATTTGCTGACCTCTGCGCACGCTCTCAAGAAAGATTTAGTCAACAGACAGACAA 532
Qy 481 GAGGAGATTAAGTGGCGGACCTCCGCGTGGCATCACCTGATCTTGAAGACAGACCTTGTG 540
Db 533 GAGGAGATTAAGTGGCGGACCTCCGCGTGGCATCACCTGATCTTGAAGACAGACCTTGTG 592
Qy 541 AAGTCAGAGCTGTACACATTTGAATCTCAAGTGGAGCTTTCAGATTTCAGATTTCAGGA 600
Db 593 AAGTCAGAGCTGTACACATTTGAATCTCAAGTGGAGCTTTCAGATTTCAGATTTCAGGA 652
Qy 601 AGAAGGATTTCTGAGGTTT 619
Db 653 AGAAGGATTTCTGAGGTTT 671

RESULT 13
CB997527 776 bp mRNA linear EST 01-MAY-2003
LOCUS AGENCOURT 13620640 NIH_MGC_148 Homo sapiens cDNA clone
DEFINITION IMAGE:30338684 5', mRNA sequence.
ACCESSION CB997527
VERSION CB997527.1 GI:30292047
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 776)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strauberg, Ph.D.
Email: egarbo-remail.nih.gov
Tissue Procurement: Dr. Stefan Hanson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: NDAM365 row: 1 column: 21
High quality sequence stop: 564.
Location/Qualifiers
1..776
```

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30338684"
/issue_type="pre-clampic placenta"
/lab_host="DH10B Tona"
/clone_idb="NIH_MGC_148"
/note="Organ: placenta; Vector: pBluescriptR; Site_1:
all-XhoI; Site_2: BamH; Library is oligo-dT primed and
directionally cloned using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert
size 2.3 kb and normalized to ROT 5. This is a primary
library enriched for full-length clones and constructed
using the Cap-trapper method (Carninci, in preparation).
Library constructed by M. Brownstein (NIH/NHGRI,
National Institutes of Health). Note: this is a NIH_MGC
Library."

ORIGIN
Query Match 22.0%; Score 461; DB 6; Length 776;
Best Local Similarity 99.5%; Pred. No. 1,8e-241;
Matches 611; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGAGAGGTTTCTGTTACTATATGCTACACAGACGGGACAGCGAAAGCCATCGACAA 60
Db 88 ATGAGAGGTTTCTGTTACTATATGCTACACAGACGGGACAGCGAAAGCCATCGACAA 147
Qy 61 GAAATGTGAGCAAGCTGTGTGATACATGATATTTCTGCAGATCTTCACTATATTAGTAA 120
Db 148 GAAATGTGAGCAAGCTGTGTGATACATGATATTTCTGCAGATCTTCACTATATTAGTAA 207
Qy 121 TCCGATTAAGTATGACCTTAACCAACCGAAACGCTCTCTTGTGTGTGTTTCTACACG 180
Db 208 TCCGATTAAGTATGACCTTAACCAACCGAAACGCTCTCTTGTGTGTGTTTCTACACG 267
Qy 181 GGCACCGGAGACCCACCCGACACAGCCCGAAGTTGTTAAGGAAATACAGAACCAACA 240
Db 268 GGCACCGGAGACCCACCCGACACAGCCCGAAGTTGTTAAGGAAATACAGAACCAACA 327
Qy 241 CTGCGCGTGTATTTCTTGTCTCACTGCGGTATGAGTTAGCTGCGTCTCGGTGATTGAA 300
Db 328 CTGCGCGTGTATTTCTTGTCTCACTGCGGTATGAGTTAGCTGCGTCTCGGTGATTGAA 387
Qy 301 TACACCTACTTTTGCATTTGGGGGAGATTAATGATTAACGACTTCAAGCTTGGAGCC 360
Db 388 TACACCTACTTTTGCATTTGGGGGAGATTAATGATTAACGACTTCAAGCTTGGAGCC 447
Qy 361 CCGCATTTCTATGACATGACATGACATGACATGACATGACATGACATGACATGACATGAC 420
Db 448 CCGCATTTCTATGACATGACATGACATGACATGACATGACATGACATGACATGACATGAC 507
Qy 421 CCGTGATTTGCTGACCTCTGCGCACGCTCTCAAGAAAGATTTAGTCAACAGACAGACAA 480
Db 508 CCGTGATTTGCTGACCTCTGCGCACGCTCTCAAGAAAGATTTAGTCAACAGACAGACAA 567
Qy 481 GAGGAGATTAAGTGGCGGACCTCCGCGTGGCATCACCTGATCTTGAAGACAGACCTTGTG 540
Db 568 GAGGAGATTAAGTGGCGGACCTCCGCGTGGCATCACCTGATCTTGAAGACAGACCTTGTG 627
Qy 541 AAGTCAGAGCTGTACACATTTGAATCTCAAGTGGAGCTTTCAGATTTCAGATTTCAGGA 600
Db 628 AAGTCAGAGCTGTACACATTTGAATCTCAAGTGGAGCTTTCAGATTTCAGATTTCAGGA 687
Qy 601 AGAAGGATTTCTGA 614
Db 688 AGAAGGATTTCTGA 701

RESULT 14
AUI32586 822 bp mRNA linear EST 01-AUG-2002
LOCUS AUI32586
DEFINITION AUI32586 NT2RP4 Homo sapiens cDNA clone NT2RP4000141 5', mRNA
sequence.
```

ACCESSION	AU132586
VERSION	AU132586.1 GI:10992940
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrates; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 822) Oca,T., Sugiyama,T., Ishii,S., Suzuki,Y., Saito,K., Yamamoto,J., Nishikawa,T., Nakamura,Y., Nagai,T., Sugano,S., Masuno,Y. and Isogai,T. HRI human cDNA project (Oca,T., Sugiyama,T., Ishii,S., Suzuki,Y., Saito,K., Yamamoto,J., Nishikawa,T., Nakamura,Y., Nagai,T., Sugano,S., Masuno,Y. and Isogai,T.) Unpublished (2000) Contact: Takao Isogai Genomics Laboratory Helix Research Institute 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan Tel.: 81-438-52-3975 Fax: 81-438-52-3986 Email: genomics@hri.co.jp HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute. Location/Qualifiers 1..822 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="NT2RP4000141" /cell_type="teratocarcinoma" /cell_line="NT2" /clone_1lb="NT2RP4" /note="Vector: pMEB18FUD; mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction"
TITLE	
JOURNAL	
COMMENT	
FEATURES	
Source	
Query Match	21.7%; Score 455; DB 1; Length 822;
Best Local Similarity	99.6%; Pred. No. 3.6e-28;
Matches 555; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
DG	283 GGTCGCGTGAATTGAGATACCACTACTTTTGCAATGGGGAGAATAATTGATTAACA 342
DG	182 GGTCCGGTAGTTACAGATACCACTACTTTTGCAATGGGGAGAATAATTGATTAACA 241
DG	343 CTTCAAGAGCTTGAAGCCCGGCATTTCTATGACACTGACATGCAAGATGACTGTGTAGT 402
DG	242 CTTCAAGAGCTTGAAGCCCGGCATTTCTATGACACTGACATGCAAGATGACTGTGTAGT 301
DG	403 TTAAAGCTGTGGTGAAGCCGCTGATTTGCTGACACTGGGCCAAGCCTCAGAAACATTTT 462
DG	302 TTAAGACTGTGGTGAAGCCGCTGATTTGCTGACACTGGGCCAAGCCTCAGAAACATTTT 361
DG	463 AGGTCAAGCAGAGACAAGAGAGATAGTGGCCATCCCGGTGGCATCACTGCATCC 522
DG	362 AGGTCAAGCAGAGACAAGAGAGATAGTGGCCATCCCGGTGGCATCACTGCATCC 421
DG	523 TTGAGGACAGACCTTTGTAAGTCAAGAGCTGCTACACATTGAATTCGAAGTCAAGCTTCTG 582
DG	422 TCGAGGACAGACCTTTGTAAGTCAAGAGCTGCTACACATTGAATTCGAAGTCAAGCTTCTG 481
DG	583 AGATTGAGTGAATTGAGGAAGAAAGATTCTGAGGTTTTGAAGCAAATGACAGTAAAGC 642
DG	482 AGATTGAGTGAATTGAGGAAGAAAGATTCTGAGGTTTTGAAGCAAATGACAGTAAAGC 541
DG	643 AACCAATCCAATGTTGTAATTAAGAAGCTTGAAGTCTTCACTTACCCTTCCGTACCCCCA 702
DG	542 AACCAATCCAATGTTGTAATTAAGAAGCTTGAAGTCTTCACTTACCCTTCCGTACCCCCA 601
DG	703 CTCACAGAGCTCTCTGAATATTCCTGTGTTAACCCCAAAATTTNACAGGTACATCTG 762

Db	602	CTCTCAACAAGCCTCTCGAATATATCTCGATTACCCCAAGAAATTTACAGGTAAACATCTG	661
Qy	763	CAGAGATCTCTTGCCCGCAGAGAAAGCAGATCTGTACTTTCAGACATTCAGATTTTT	822
Db	662	CAGAGATCTCTTGCCCGCAGAGAAAGCAGATCTGTGTGACTTCAGCAGATTCAGATTTTT	721
Qy	823	CAAGTGCATTTTCAAA	839
Db	722	CAGTGCATTTTCAAA	738
RESULT 15			
LOCUS	AM965709	591 bp	mRNA linear. EST 01-JUN-2000
DEFINITION	EST377782	WAGE resequences, MAGI	Homo sapiens cDNA, mRNA sequence.
ACCESSION	AM965709		
VERSION	AM965709.1	GI:8155545	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Ekharjola; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Hogde, P., Qi, R., Abernathy, K., Dharap, S., Gaupard, R., Gay, C., Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and Quackenbush, J.		
TITLE	Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: John Quackenbush The Institute for Genomic Research 9112 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 3528 Fax: 301 838 0208 Email: johng@tigr.org plate: 218		
FEATURES	Seq primer: Reverse.		
source	Location/Qualifiers		
	1..591		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone_id="WAGE resequences, MAGI"		
	/note="Vector: Bluescriptskm"		
ORIGIN			
Query Match	21.4%; Score 448; DB 2; Length 591;		
Best Local Similarity	100.0%; Pied. No. 2.5e-234; Indels 0; Gaps 0;		
Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1192	AAGCGCAGGCTACAGAGCTGTGCAGTAAACAAGGGGACCCGATTTAGCCGCTTTGTA	1251
Db	1	AAGCGCAGGCTACAGAGCTGTGCAGTAAACAAGGGGACCCGATTTAGCCGCTTTGTA	60
Qy	1252	CGAGATGCTGTGCTGCTTGTGGATCTGCTCCGCTTCCCTTCTTCCGACGACCA	1311
Db	61	CGAGATGCTGTGCTGCTTGTGGATCTGCTCCGCTTCCCTTCTTCCGACGACCA	120
Qy	1312	CTCAGTCTCTGTGCTGCAAACTTCTTCTTAACTTCAACCCGACCATATTGTGTGCAAGC	1371
Db	121	CTCAGTCTCTGTGCTGCAAACTTCTTCTTAACTTCAACCCGACCATATTGTGTGCAAGC	180
Qy	1372	TCAAGTTATTTTCAACGAGAAAGCTCATTTTGTCTTCAACATTTGTGGAATTTCTGCT	1431
Db	181	TCAAGTTATTTTCAACGAGAAAGCTCATTTTGTCTTCAACATTTGTGGAATTTCTGCT	240
Qy	1432	ACTGCACACAAGAGGTTCTGCGAAGGAGATATGACAGGCTGCTGCTGCTGTTGTT	1491
Db	241	ACTGCACACAAGAGGTTCTGCGAAGGAGATATGACAGGCTGCTGCTGCTGTTGTT	300
Qy	1492	GCTTCAGTTCTTTCAGCCAACTATTCATGCAATCCCATGAAACAAGCCGGAAAGCCCTGGCT	1551
Db	301	GCTTCAGTTCTTTCAGCCAACTATTCATGCAATCCCATGAAACAAGCCGGAAAGCCCTGGCT	360

Qy 1552 CCTAAGATATCCATCTCTCTCGAACAATAATTTCTTCCACTTACCAATGACCCCTCA 1611  
Db 361 CCTAAGATATCCATCTCTCTCGAACAATAATTTCTTCCACTTACCAATGACCCCTCA 420  
Qy 1612 ATCCCATCATATATGATGGTCCAGGA 1639  
Db 421 ATCCCATCATATATGATGGTCCAGGA 448

Search completed: August 27, 2005, 15:58:35  
Job time : 4548.04 secs





|||||  
Db 320 CTGCGGTTATTTCTTGTCTCACTGCGGTATGGTTACTGGGTCTCGGTATTCAGAA 379  
|||  
Qy 301 TACACCTACTTTTTCAGATGGGGGAGATATGATTAAGCACTTCAAGAGCTTGGAGCC 360  
|||  
Db 380 TACACCTACTTTTTCAGATGGGGGAGATATGATTAAGCACTTCAAGAGCTTGGAGCC 439  
|||  
Qy 361 CGGCAATTTCTATGACACTGACATGACATGACTGTGTAGGTTTAGAACTTGTGTGAG 420  
|||  
Db 440 CGGCAATTTCTATGACACTGACATGACATGACTGTGTAGGTTTAGAACTTGTGTGAG 499  
|||  
Qy 421 CCGTGATTTGCTGACCTCTGGCCAGCCCTCAGAAAGATTTTATAGTAAAGCAGAGCA 480  
|||  
Db 500 CCGTGATTTGCTGACCTCTGGCCAGCCCTCAGAAAGATTTTATAGTAAAGCAGAGCA 559  
|||  
Qy 481 GAGGAGATTAAGTGGCGGCACTCCGGTGGCATCACTGCTCTTGAAGAGAGAGCTTGTG 540  
|||  
Db 560 GAGGAGATTAAGTGGCGGCACTCCGGTGGCATCACTGCTCTTGAAGAGAGAGCTTGTG 619  
|||  
Qy 541 AAGTCAGAGCTGCTACACATTTGAATCTCAAGTCAGAGCTTCTGAGATTCAGATTCAGA 600  
|||  
Db 620 AAGTCAGAGCTGCTACACATTTGAATCTCAAGTCAGAGCTTCTGAGATTCAGATTCAGA 679  
|||  
Qy 601 AGAAGGATTTCTGAGCTTTTGAAGCAAAATGCAAGTGAACGAAACCAATCCATTTGTA 660  
|||  
Db 680 AGAAGGATTTCTGAGCTTTTGAAGCAAAATGCAAGTGAACGAAACCAATCCATTTGTA 739  
|||  
Qy 661 ATTGAAGATTTGAGCTCTCACTTACCGGTTCCGTTACCCCACTCTCAAGAGCTCTCTG 720  
|||  
Db 740 ATTGAAGATTTGAGCTCTCACTTACCGGTTCCGTTACCCCACTCTCAAGAGCTCTCTG 799  
|||  
Qy 721 AATATCTGCTGTTTACCCCGAATATTTTACAGTATCTGACAGAGCTCTTGGCCAG 780  
|||  
Db 800 AATATCTGCTGTTTACCCCGAATATTTTACAGTATCTGACAGAGCTCTTGGCCAG 859  
|||  
Qy 781 GAGGAAGCCAGATATCTGTGACTTTCAGAGATCTCACTTTTTCAGTCCAAATTTCAAG 840  
|||  
Db 860 GAGGAAGCCAGATATCTGTGACTTTCAGAGATCTCACTTTTTCAGTCCAAATTTCAAG 919  
|||  
Qy 841 GCAGTTCACTTACCTAAGATGATGATCAATTAAGCACTCTGCTGTAGAAATTTGACAT 900  
|||  
Db 920 GCAGTTCACTTACCTAAGATGATGATCAATTAAGCACTCTGCTGTAGAAATTTGACAT 979  
|||  
Qy 901 TCAAAATCAAGCTTTTCTATCAGCTGTGAGATGCTTCAAGCTGATCTGCTTCAACAT 960  
|||  
Db 980 TCAAAATCAAGCTTTTCTATCAGCTGTGAGATGCTTCAAGCTGATCTGCTTCAACAT 1039  
|||  
Qy 961 GATTCAGAGTACAAAGCTTCTCAAGAGCTGAGCTTGAAGTAAAGAGAGCACTGC 1020  
|||  
Db 1040 GATTCAGAGTACAAAGCTTCTCAAGAGCTGAGCTTGAAGTAAAGAGAGCACTGC 1099  
|||  
Qy 1021 GTCCTTTGAAAATTAAGGAGAGCACAAAAGAAAGAAAGAGTACCTTACCCAGCATATA 1080  
|||  
Db 1100 GTCCTTTGAAAATTAAGGAGAGAGCACAAAAGAAAGAAAGAGTACCTTACCCAGCATATA 1159  
|||  
Qy 1081 CTGCGGAGATGTTCTCTCAGATTCATTTTACCTGCTGTCTTGAATCCAGCAATTTCT 1140  
|||  
Db 1160 CTGCGGAGATGTTCTCTCAGATTCATTTTACCTGCTGTCTTGAATCCAGCAATTTCT 1219  
|||  
Qy 1141 AAAAAGGCAATTTTGGAGAGCCCTGTGAGCATATCCAGTACAGTGTGAAAAGCCGAG 1200  
|||  
Db 1220 AAAAAGGCAATTTTGGAGAGCCCTGTGAGCATATCCAGTACAGTGTGAAAAGCCGAG 1279  
|||  
Qy 1201 CTACAGAGCTGTGACAGTAAACAGAGGGGAGCCGATTAAGCCGCTTTGTACAGATGCC 1260  
|||  
Db 1280 CTACAGAGCTGTGACAGTAAACAGAGGGGAGCCGATTAAGCCGCTTTGTACAGATGCC 1339  
|||  
Qy 1261 TGTGCTGCTTGTGAGATCTCTCTGCTTCTGCTTCTTGTGACAGCACTCAAGTCTC 1320  
|||  
Db 1340 TGTGCTGCTTGTGAGATCTCTCTGCTTCTGCTTCTTGTGACAGCACTCAAGTCTC 1399  
|||  
Qy 1321 CTGCTGAAACATCTTCTTAAACCTTCAACCCAGAGCATATTCGTGTGAAAGTCAAGTTTA 1380  
|||

Db 1400 CTGCTGAAACATCTTCTTAAACCTTCAACCCAGAGCATATTCGTGTGAAAGTCAAGTTTA 1459  
|||  
Qy 1381 TTTCAACCCAGAAAGCTCCATTTTGTCTTCAACATTTGAGAAATTTCTCTACCTGCACAA 1440  
|||  
Db 1460 TTTCAACCCAGAAAGCTCCATTTTGTCTTCAACATTTGAGAAATTTCTCTACCTGCACAA 1519  
|||  
Qy 1441 ACAGAGTTCTGCGGAGAGAGATATGTACAGCTGCTGCTGCTTGTGTGCTTCAATT 1500  
|||  
Db 1520 ACAGAGTTCTGCGGAGAGAGATATGTACAGCTGCTGCTGCTTGTGTGCTTCAATT 1579  
|||  
Qy 1501 CTTGACGCAAAACATATGATCCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATTA 1560  
|||  
Db 1580 CTTGACGCAAAACATATGATCCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATTA 1639  
|||  
Qy 1561 TCCATCTCTCTGCAACCAAAATTTCTTCACTTACAGATGAGAGAGAGAGAGAGAGAGATC 1620  
|||  
Db 1640 TCCATCTCTCTGCAACCAAAATTTCTTCACTTACAGATGAGAGAGAGAGAGAGAGAGATC 1699  
|||  
Qy 1621 ATATAGTGGGTTCAGAGAACCGGATAGCCCGCTTATTTGGGTTCTTACAACTAGAGAG 1680  
|||  
Db 1700 ATATAGTGGGTTCAGAGAACCGGATAGCCCGCTTATTTGGGTTCTTACAACTAGAGAG 1759  
|||  
Qy 1681 AAATCTCAAGAAACAAACCCAGATGAGAAATTTTGAAGCAATGTTGTTTGTGCTGC 1740  
|||  
Db 1760 AAATCTCAAGAAACAAACCCAGATGAGAAATTTTGAAGCAATGTTGTTTGTGCTGC 1819  
|||  
Qy 1741 AGGATTAAGGATTAAGGATTAATCTATTCAGAAAGAGCTCAGACATTTCTTAAAGATGG 1800  
|||  
Db 1820 AGGATTAAGGATTAAGGATTAATCTATTCAGAAAGAGCTCAGACATTTCTTAAAGATGG 1879  
|||  
Qy 1801 ATCTTAATCTATTAAGGTTTCTTCTCAAGAGATGCTCTGTGTGGAGAGAGAGAGAGCC 1860  
|||  
Db 1880 ATCTTAATCTATTAAGGTTTCTTCTCAAGAGATGCTCTGTGTGGAGAGAGAGAGAGCC 1939  
|||  
Qy 1861 CCAGCAAGATTAATCAAGCAACATCCAGCTTCAAGGAGAGAGAGAGAGAGATCTCTC 1920  
|||  
Db 1940 CCAGCAAGATTAATCAAGCAACATCCAGCTTCAAGGAGAGAGAGAGAGAGATCTCTC 1999  
|||  
Qy 1921 CTCACAGAGAGAGAGAGAGATTTATGTGTGTGAGATGCAAGAAATATGAGCAAGATGTA 1980  
|||  
Db 2000 CTCACAGAGAGAGAGAGAGATTTATGTGTGTGAGATGCAAGAAATATGAGCAAGATGTA 2059  
|||  
Qy 1981 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGATG 2040  
|||  
Db 2060 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGATG 2119  
|||  
Qy 2041 AAAACCTGGCCACTTTTAAAGAAAGAAACGCTTACCTTCAAGATATTTGTGATTA 2097  
|||  
Db 2120 AAAACCTGGCCACTTTTAAAGAAAGAAACGCTTACCTTCAAGATATTTGTGATTA 2176  
|||

RESULT 2  
AX050463 3259 bp DNA linear PAT 12-JAN-2001  
LOCUS AX050463  
DEFINITION Sequence 23 from Patent WO0071754.  
ACCESSION AX050463.1 GI:12226668  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
AUTHORS Johnson, W.G., and Stenroos, R.S.  
TITLE Methods for diagnosing, preventing, and treating developmental  
disorders due to a combination of genetic and environmental factors  
JOURNAL Patent: WO 0071754-A 23 NOV-2000;  
University of Medicine and Dentistry of New Jersey (US)  
FEATURES  
source 1.3259  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

## ORIGIN

97.6% Score 2046; DB 6; Length 3259;

Query Match 100.0%; Freq. No. 0;  
Best Local Similarity 0; Mismatches 1; Indels 0; Gaps 0;  
Matches 2096; Conservative 0;

1 ATGAGAGAGTTCTGTACTATATGCTACAGCAGGAGCAGGCAAGGCCATCCGAGAA 60  
80 ATGAGAGAGTTCTGTACTATATGCTACAGCAGGAGCAGGCAAGGCCATCCGAGAA 139  
61 GAAATGTGTAGCAAGCTGTGTACATGATTTCTGCAATCTTCACTAATTAAGTAA 120  
140 GAAATGTGTAGCAAGCTGTGTACATGATTTCTGCAATCTTCACTAATTAAGTAA 199  
121 TCCGATTAAGTAACTTAATAAACCAGAACGCTCTCTGTGTGTGTGTGTGTGTGTGT 180  
200 TCCGATTAAGTAACTTAATAAACCAGAACGCTCTCTGTGTGTGTGTGTGTGTGTGT 259  
181 GGCACCGAGAGCCACCCGACACAGCCCGCAAGTTGTTAAGAAATACAGAACCAACA 240  
260 GGCACCGAGAGCCACCCGACACAGCCCGCAAGTTGTTAAGAAATACAGAACCAACA 319  
241 CTGCGGTGTGATTTCTTGTCTCACTGCGGTATGGGTATCTGGGTCTCGGTATTCAGAA 300  
320 CTGCGGTGTGATTTCTTGTCTCACTGCGGTATGGGTATCTGGGTCTCGGTATTCAGAA 379  
301 TACACTTACTTTTGCATATGGGGGAAATTAATTAATTAATTAATTAATTAATTAATTA 360  
380 TACACTTACTTTTGCATATGGGGGAAATTAATTAATTAATTAATTAATTAATTAATTA 439  
361 CGGCAATTTCTATGACACTGACATGACATGACATGACATGACATGACATGACATGACAT 420  
440 CGGCAATTTCTATGACACTGACATGACATGACATGACATGACATGACATGACATGACAT 499  
421 CCGTGATTTGCTGACCTCTGCGCCAGCCCTCAGAAACATTTTAAAGTCAAGCAGAGACA 480  
500 CCGTGATTTGCTGACCTCTGCGCCAGCCCTCAGAAACATTTTAAAGTCAAGCAGAGACA 559  
481 GAGAGATTAAGTGGCGCATCTCCGGTGGCATCACTGCACTCTTGAAGACAGACTTTGG 540  
560 GAGAGATTAAGTGGCGCATCTCCGGTGGCATCACTGCACTCTTGAAGACAGACTTTGG 619  
541 AAGTCAGACTGTGACATGATGAATCTCAAGTGAAGTCTGATGATTCAGATTCAGAG 600  
620 AAGTCAGACTGTGACATGATGAATCTCAAGTGAAGTCTGATGATTCAGATTCAGAG 679  
601 AGAAGAGATTTCTGAGGTTTGAAGCAAAATGACAGTAAAGCAACCAATCCATGTTGTA 660  
680 AGAAGAGATTTCTGAGGTTTGAAGCAAAATGACAGTAAAGCAACCAATCCATGTTGTA 739  
661 ATTGAAGACTTTGAGTCTCTCACTTACCCGTTGGTACCCCACTCTCAAGCCTCTCTG 720  
740 ATTGAAGACTTTGAGTCTCTCACTTACCCGTTGGTACCCCACTCTCAAGCCTCTCTG 799  
721 AATATTCCTGTTTACCCCGCAAAATTAATTAAGTAACTGACAGAGAGTCTTGGCCAG 780  
800 AATATTCCTGTTTACCCCGCAAAATTAATTAAGTAACTGACAGAGAGTCTTGGCCAG 859  
781 GAGAGAGCCAGAGTATCTGTGACTTCAAGCAGATCCAGTCTTCAAGTCCCAATTTCAAG 840  
860 GAGAGAGCCAGAGTATCTGTGACTTCAAGCAGATCCAGTCTTCAAGTCCCAATTTCAAG 919  
841 GCAGTTCAACTTACGAAATGATGCAATTAATAAACAATCTCTGTGTGTGTGTGTGTGT 900  
920 GCAGTTCAACTTACGAAATGATGCAATTAATAAACAATCTCTGTGTGTGTGTGTGTGT 979  
901 TCAATTAACACATTTTCTATCAAGCCTGAGAGAGTCTTCAAGCAGTCCGCTTAACAGT 960  
980 TCAATTAACACATTTTCTATCAAGCCTGAGAGAGTCTTCAAGCAGTCCGCTTAACAGT 1039  
961 GATTTGAGGTACAAAGCCTTACTCAAGAGCTGAGAGTCTTCAAGTAAAGAGAGACTGC 1020  
1040 GATTTGAGGTACAAAGCCTTACTCAAGAGCTGAGAGTCTTCAAGTAAAGAGAGACTGC 1099

1021 GTCTTTTGAATAAAGCAGACACAAGAGAGAAAGAGCTTACCTTACCCGAGATATA 1080  
1100 GTCTTTTGAATAAAGCAGACACAAGAGAGAAAGAGCTTACCTTACCCGAGATATA 1159  
1081 CCGCGGAGATGTTCTCTCAAGTTCATTTTACCTGTGTCTTGAATAATCCGAGCAATTCCT 1140  
1160 CCGCGGAGATGTTCTCTCAAGTTCATTTTACCTGTGTCTTGAATAATCCGAGCAATTCCT 1219  
1141 AAAAGGAGATTTTGGAGCCCTGTGAGCTATACAGTGAACAGTGTGAAAAGGCGAGG 1200  
1220 AAAAGGAGATTTTGGAGCCCTGTGAGCTATACAGTGAACAGTGTGAAAAGGCGAGG 1279  
1201 CTACAGAGCTGTGAGCTAACAAGGGGAGCCGATTAAGCCGCTTGTACAGAGATGCC 1260  
1280 CTACAGAGCTGTGAGCTAACAAGGGGAGCCGATTAAGCCGCTTGTACAGAGATGCC 1339  
1261 TGTGCTGT 1320  
1340 TGTGCTGT 1399  
1321 CTGCTGGAACATCTTCTTAACTTGAACCAACCAAGCAATTAAGTGTGCAAGCTCAAGTTA 1380  
1400 CTGCTGGAACATCTTCTTAACTTGAACCAACCAAGCAATTAAGTGTGCAAGCTCAAGTTA 1459  
1381 TTTCAACCCAGAAAGCTCAATTTGTCTTCAACATTTGTGAATTTCTGTCTACCTGCA 1440  
1460 TTTCAACCCAGAAAGCTCAATTTGTCTTCAACATTTGTGAATTTCTGTCTACCTGCA 1519  
1441 ACAGAGTTCTGCGGAGGAGATGATACAGCTGCGTGTGTGTGTGTGTGTGTGTGTGTGT 1500  
1520 ACAGAGTTCTGCGGAGGAGATGATACAGCTGCGTGTGTGTGTGTGTGTGTGTGTGTGT 1579  
1501 CTTGAGCCAAATATATGATATCCCATGAAGCAACCGGAAAGCCCTGCTCTTAAGATA 1560  
1580 CTTGAGCCAAATATATGATATCCCATGAAGCAACCGGAAAGCCCTGCTCTTAAGATA 1639  
1561 TCCATCTCTCTGGAACAACAATTTCTTCACTTACAGATGACCCCTCAATGCCCATC 1620  
1640 TCCATCTCTCTGGAACAACAATTTCTTCACTTACAGATGACCCCTCAATGCCCATC 1699  
1621 ATATATGT 1680  
1700 ATATATGT 1759  
1681 AAACTCCAAAGAACCAACCAAGTGAATTTTGAAGCAATGTGTGTGTGTGTGTGTGTGT 1740  
1760 AAACTCCAAAGAACCAACCAAGTGAATTTTGAAGCAATGTGTGTGTGTGTGTGTGTGT 1819  
1741 AGGCAATTAAGGATTAAGGATTAATTAATTAAGAAAGCTGACATTTCTTAAAGCATGG 1800  
1820 AGGCAATTAAGGATTAAGGATTAATTAATTAAGAAAGCTGACATTTCTTAAAGCATGG 1879  
1801 ATCTTAATCTCATTAAGGATTTCTTCTCAAGAGATCTCTGTGTGTGTGTGTGTGTGTGT 1860  
1880 ATCTTAATCTCATTAAGGATTTCTTCTCAAGAGATCTCTGTGTGTGTGTGTGTGTGTGT 1939  
1861 CCAAGCAAGTATGTACAGACCAATCCAGCTTCAATGCGCAAGAGTGTGCGAGAAATCTCTC 1920  
1940 CCAAGCAAGTATGTACAGACCAATCCAGCTTCAATGCGCAAGAGTGTGCGAGAAATCTCTC 1999  
1921 CTCGAGAGAAAGGCGCATTTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1980  
2000 CTCGAGAGAAAGGCGCATTTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2059  
1981 CATGATGCCCTTGTGCAATTAATTAAGCAAGGTTGTGAGTTGAAAACCTAGAAAGCAATG 2040  
2060 CATGATGCCCTTGTGCAATTAATTAAGCAAGGTTGTGAGTTGAAAACCTAGAAAGCAATG 2119  
2041 AAAACCTGTGCACTTTTAAAGAAAGAAAGGCTACCTTCAAGATTTTGTGTCAATA 2097  
2120 AAAACCTGTGCACTTTTAAAGAAAGAAAGGCTACCTTCAAGATTTTGTGTCAATA 2176

RESULT 3  
LOCUS AF025794 3259 bp mRNA linear PRI 26-MAR-1998  
DEFINITION Homo sapiens methionine synthase reductase (MTRR) mRNA, complete cds.  
ACCESSION AF025794  
VERSION AF025794.1 GI:2981302  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 3259)  
AUTHORS Leclerc,D., Wilson,A., Dumas,R., Gafnick,C., Song,D., Watkins,D., Heng,H.H.O., Rommens,U.M., Scherer,S.W., Rosenblatt,D.S. and Gravel,R.A.  
TITLE Cloning and mapping of a cDNA for methionine synthase reductase, a flavoprotein defective in patients with homocystinuria  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (6), 3059-3064 (1998)  
MEDLINE 98169496  
PubMed 9501215  
REFERENCE 2 (bases 1 to 3259)  
AUTHORS Leclerc,D.  
TITLE Direct Submission  
JOURNAL Submitted (19-SEP-1997) Human Genetics, McGill University - Montreal Children's Hospital Research Institute, 4060 Ste-Catherine West, Montreal, Que H3Z 2Z3, Canada  
REFERENCE 3 (bases 1 to 3259)  
AUTHORS Leclerc,D.  
TITLE Direct Submission  
JOURNAL Submitted (12-NOV-1997) Human Genetics, McGill University - Montreal Children's Hospital Research Institute, 4060 Ste-Catherine West, Montreal, Que H3Z 2Z3, Canada  
REMARK Sequence update by submitter

FEATURES  
source location/Qualifiers  
1..3259  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/map="Sp15.2-p15.3"  
1..3259  
/gene="MTRR"  
80..2176  
/gene="MTRR"  
/note="B12 vitamin; cblB complementation group"  
/codon\_start=1  
/product="methionine synthase reductase"  
/db\_xref="GI:2981303"  
/protein\_id="AAC39667.1"  
translation="MRRGLLDYIAQQGAKRAIEMGEQAVVHFPSADIIICISDEKYDKETRIPLVVVSTGTGGDDPPDPAKFKVEIKTQTLPDFPALKIRIGLIGDSERTTYFCNGKGITDKRELDELGAHHYYDTGHADDGVLEIVPEWTAIGLMPILRRFPSSRQBEIGALFVPASIPSLRTDLVNSIELIHLESQVELRPDSDGRKOSEVLKONAVNSQSNVVIDPFESSLTRSVPLSPQSASINLPGAPLEYLVQHLOESLGQESVSOTSAVPVFQPIISAVALQTTMDAIKTLLVLELDINSNDFSPGPDARSVCIPNSDSVOGLORLOLEDKRBHVCLTKAKDTKKKGATLPQHTIAGCSIOIFPMCKEITRAPILPKAATRLAYVTSDSARKRRLQELSCKQCAADVSRVVRDACILLIILAAPSOPPSILIEHLKPQRBPYSICASSLPHGISISPRITNSFHLDPDPSIPINVGPGTGIAPIGFIQHRKIQEOEHHEHGSKALSFHSISPTNTNFHLPDPSIPINVGPGTGIAPIGFIQHRKIQEOEH PDGNFGMMILFFGCGRHKRDYLFRKELRHFLKHGILTTLTKVTSFSRDA PVGEERAPACYVDNNQLHGQOOVARILLQDNENHIYVCGDAKNMAVDVHALVQIISKEVGEKLEAMKT LATLKEERYTIQDIWS"

ORIGIN  
Query Match 97.6%; Score 2046; DB 9; Length 3259;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2096; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 1 ATGAGAGAGTTCTTCTATATGCTACACAGCAGGAGCAGGCAGAAAGCATTCGCCGAA 60  
ATTGAGAGAGTTCTTCTATATGCTACACAGCAGGAGCAGGCAGAAAGCCATGCGCAA 139

Oy	61	GAATATGCTAGACGACGTGCTGCTACATGATATTTCTGCAGATCTCTCACTAATATATGAA	120
Dp	140	GAATATGCTAGACGACGTGCTGCTACATGATATTTCTGCAGATCTCTCACTAATATGAA	199
Oy	121	TCCGATATAGTATGACTTAAACCCGAAACAGCTCCTCTGTTGTGTGTCTTACACAG	180
Dp	200	TCCGATATAGTATGACTTAAACCCGAAACAGCTCCTCTGTTGTGTGTCTTACACAG	259
Oy	181	GGCACCCGAGACCCACCCGACACAGCCCGCAATTTGTATAGAAATATACAAACCAACA	240
Dp	260	GGCACCCGAGACCCACCCGACACAGCCCGCAATTTGTATAGAAATATACAAACCAACA	319
Oy	241	CTGCACGGTGTATTTCTTTGTCTCACTGCGGTATAGGTATCTCGGTATTCAGAA	300
Dp	320	CTGCACGGTGTATTTCTTTGTCTCACTGCGGTATAGGTATCTCGGTATTCAGAA	379
Oy	301	TACACCTACTTTTGCAATGCGGGGAGATATATGATAAACGACTTCAAGACTTGGAGCC	360
Dp	380	TACACCTACTTTTGCAATGCGGGGAGATATATGATAAACGACTTCAAGACTTGGAGCC	439
Oy	361	CGGATTTCTATGACACTGACATGACAGATGACGTGTAGGTATAGAACTTGTGTGTAG	420
Dp	440	CGGATTTCTATGACACTGACATGACAGATGACGTGTAGGTATAGAACTTGTGTGTAG	499
Oy	421	CCGTGATTTGCTGACCTCTGGCCAGCCCTCAGAAAGACTTTTAGCTCAAGCAGACGACA	480
Dp	500	CCGTGATTTGCTGACCTCTGGCCAGCCCTCAGAAAGACTTTTAGCTCAAGCAGACGACA	559
Oy	481	GAGAGATTAAGTGGCGCACTCCCGGTGGCATCACCTGCATCTCTTGAGACAGACCTTGTG	540
Dp	560	GAGAGATTAAGTGGCGCACTCCCGGTGGCATCACCTGCATCTCTTGAGACAGACCTTGTG	619
Oy	541	AAGTCAGAGCTGCATACATTTGATCTCAAGTCAGCTTCTGAGATTCGATGATTCAGAA	600
Dp	620	AAGTCAGAGCTGCATACATTTGATCTCAAGTCAGCTTCTGAGATTCGATGATTCAGAA	679
Oy	601	AGAAAGATTTCTGAGTTTGTGAGCAAAATGCAGTGAAACAGCAACCAATCCATGTGTGA	660
Dp	680	AGAAAGATTTCTGAGTTTGTGAGCAAAATGCAGTGAAACAGCAACCAATCCATGTGTGA	739
Oy	661	ATTGAAACATTTGAGTCTCTCACTTAAACCGTTGGTACCCCCACTCTCAAGACCTCTCTG	720
Dp	740	ATTGAAACATTTGAGTCTCTCACTTAAACCGTTGGTACCCCCACTCTCAAGACCTCTCTG	799
Oy	721	AATATTCCTGTTTAAACCCCGAGATATTTACAGGTACATCTGCAGAGTCTTGGCCAG	780
Dp	800	AATATTCCTGTTTAAACCCCGAGATATTTACAGGTACATCTGCAGAGTCTTGGCCAG	859
Oy	781	GAGGAAAGCCAGATATCTGTGACTTTCAGCAGATCCAGTTTTCAAAGTCCCAATTTCAAG	840
Dp	860	GAGGAAAGCCAGATATCTGTGACTTTCAGCAGATCCAGTTTTCAAAGTCCCAATTTCAAG	919
Oy	841	GCAGTTCAACTTACATGAAATGATGSCATATAAAACACTCTGCTGTATGAAATTTGACACTT	900
Dp	920	GCAGTTCAACTTACATGAAATGATGSCATATAAAACACTCTGCTGTATGAAATTTGACACTT	979
Oy	901	TCAATATCAGACTTTTCCATCAGCTTGAGATGCTTCAAGCTGATCTGCCCTPAACAGT	960
Dp	980	TCAATATCAGACTTTTCCATCAGCTTGAGATGCTTCAAGCTGATCTGCCCTPAACAGT	1039
Oy	961	GATTTCAAGGTACAAAGCTTACCTCAAAAGCTGAGCTGAAAGTATAAAGAGACACATGCG	1020
Dp	1040	GATTTCAAGGTACAAAGCTTACCTCAAAAGCTGAGCTGAAAGTATAAAGAGACACATGCG	1099
Oy	1021	GTCCCTTTTGAATAATAAAGCAGACACAAAGAGAAAGAGCTTACCCCAACATATA	1080
Dp	1100	GTCCCTTTTGAATAATAAAGCAGACACAAAGAGAAAGAGCTTACCCCAACATATA	1159
Oy	1081	CTGCGGGAGTCTCTTCACAGTCAATTTTACCTGTGTCTTGAATTCGAGCAATTTCTT	1140
Dp	1160	CTGCGGGAGTCTCTTCACAGTCAATTTTACCTGTGTCTTGAATTCGAGCAATTTCTT	1219

```

Oy 1141 AAAAGGCAATTTTGGGAGCCCTTGTGACATATACAGTAGAGTGTGAAAAAGCCAG 1200
Db 1220 AAAAGGCAATTTTGGGAGCCCTTGTGACATATACAGTAGAGTGTGAAAAAGCCAG 1279
Oy 1204 CTACAGAGCTGTGACATTAACAGAGGAGCCGATTAATAGCCGCTTTGTACAGATGCG 1260
Db 1280 CTACAGAGCTGTGACATTAACAGAGGAGCCGATTAATAGCCGCTTTGTACAGATGCG 1339
Oy 1261 TGTGCTGTGTGTGATCTCCCTCCGCTTCCCTTCCCTTCCGACAGCACACAGCTC 1320
Db 1340 TGTGCTGTGTGTGATCTCCCTCCGCTTCCCTTCCCTTCCGACAGCACACAGCTC 1399
Oy 1321 CTGCTGAAATCTTCTTAACTTCAACCCAGACATATTCGTGTGCAAGCTCAAGTTA 1380
Db 1400 CTGCTGAAATCTTCTTAACTTCAACCCAGACATATTCGTGTGCAAGCTCAAGTTA 1459
Oy 1381 TTTCACCCAGAAAGCTCAATTTGTCTTCAACATTTGTGAAATTTCTGTCTACTGCGACA 1440
Db 1460 TTTCACCCAGAAAGCTCAATTTGTCTTCAACATTTGTGAAATTTCTGTCTACTGCGACA 1519
Oy 1441 AAGAGGTTCTGCGAGAGGAGATATACAGCTGCTGCTGCTGTGTGTGCTTACATT 1500
Db 1520 AAGAGGTTCTGCGAGAGGAGATATACAGCTGCTGCTGCTGTGTGTGCTTACATT 1579
Oy 1501 CTTCAGCCAAACATACATGATCCATGAGAGACAGGGGAAAGCCCTGCTCTTAAGATA 1560
Db 1580 CTTCAGCCAAACATACATGATCCATGAGAGACAGGGGAAAGCCCTGCTCTTAAGATA 1639
Oy 1561 TCATCTCTCTCGAACAAATTTCTTCACTTCAACAGTAGACCCCTCAATCCCATC 1620
Db 1640 TCATCTCTCTCGAACAAATTTCTTCACTTCAACAGTAGACCCCTCAATCCCATC 1699
Oy 1621 ATATATGTGTGTGACAGAAACCCGACATAGCCCTTATTTGGGTCTTACACATTAAGAG 1680
Db 1700 ATATATGTGTGTGACAGAAACCCGACATAGCCCTTATTTGGGTCTTACACATTAAGAG 1759
Oy 1681 AAATCTCAAGAAACAACCCAGATGAGAAATTTTGAGACATGTGTGTTTTTGCTGC 1740
Db 1760 AAATCTCAAGAAACAACCCAGATGAGAAATTTTGAGACATGTGTGTTTTTGCTGC 1819
Oy 1741 AAGCATTAAGATAGGATTAATCTATTCAGAAAAGACTCAGACATTTCTTAAGATGAG 1800
Db 1820 AAGCATTAAGATAGGATTAATCTATTCAGAAAAGACTCAGACATTTCTTAAGATGAG 1879
Oy 1801 ATCTTAATCACTTAAGAGTTTCTTCTCAAGAGATGCTCTGTGTGGGAGAGAAAGCC 1860
Db 1880 ATCTTAATCACTTAAGAGTTTCTTCTCAAGAGATGCTCTGTGTGGGAGAGAAAGCC 1939
Oy 1861 CCAGCAAGATATGTAACAAGCAATCCAGCTTCATGGCCAGAGGTGGGAGAAATCCTC 1920
Db 1940 CCAGCAAGATATGTAACAAGCAATCCAGCTTCATGGCCAGAGGTGGGAGAAATCCTC 1999
Oy 1921 CTTCAGAGAGAGCGGCATATTTATGTGTGTGAGATGCAAGAAATATGGCCAGAGATGTA 1980
Db 2000 CTTCAGAGAGAGCGGCATATTTATGTGTGTGAGATGCAAGAAATATGGCCAGAGATGTA 2059
Oy 1981 CATGATGCTCTGTGCAAAATATATAGCAAAAGAGTTGAGAGTTGAAAACTAGAGCAATG 2040
Db 2060 CATGATGCTCTGTGCAAAATATATAGCAAAAGAGTTGAGAGTTGAAAACTAGAGCAATG 2119
Oy 2041 AAAACCTGAGCCACTTAAGAAAGAAAGCGTACCTTCAGAGATATTTGGTCATTA 2097
Db 2120 AAAACCTGAGCCACTTAAGAAAGAAAGCGTACCTTCAGAGATATTTGGTCATTA 2176

```

```

RESULT 4
AF121214 3291 bp mRNA linear PRI 12-DEC-1999
LOCUS Homo sapiens methionine synthase reductase (MTRR) mRNA, complete
DEFINITION CDS.
ACCESSION AF121214
VERSION AF121214.1 GI:6561338
KEYWORDS

```

```

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 3291)
AUTHORS Leclerc,D., Odievre,M., Wu,Q., Wilson,A., Huizenga,J.J., Rozen,R., Scherer,S.W. and Gravel,R.A.
TITLE Molecular cloning, expression and physical mapping of the human methionine synthase reductase gene
JOURNAL Gene 240 (1), 75-88 (1999)
MEDLINE 20033550
PUBMED 10564814
REFERENCE 2 (bases 1 to 3291)
AUTHORS Leclerc,D., Odievre,M.-H., Wu,Q., Wilson,A., Huizenga,J.J., Johns,T., Shoubiridge,E.A., Rosenblatt,D.S., Scherer,S.W., Rozen,R. and Gravel,R.A.
TITLE Direct Submission
JOURNAL Submitted (18-JAN-1999) Human Genetics, Montreal Children's Hospital, 4060 Ste-Catherine West, Montreal, Quebec H3Z 2Z3, Canada
FEATURES
source
1..3291
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="5"
/map="5p15.2-p15.3; between W11755 and D5S1957"
1..3291
/gene="MTRR"
1..30
/gene="MTRR"
/evidence=not_experimental
31..2208
/gene="MTRR"
/EC_number="2.1.1.135"
/function="maintenance of methionine synthase in an active state"
/note="alternatively spliced; alternative splicing results in extended 5' sequence compared to GenBank Accession Number AF025794; B12 vitamin; cblE complementation group"
/codon_start=1
/product="methionine synthase reductase"
/protein_id="AAPI6876.1"
/db_xref="GI:6561339"
/translation="MGASVPRGAGLIVRYALCSPTVTCLEWVRPFLIATAGCGAKA
IAEMCEQAVHGFSDADLHCSISPKYDKETATPLVYVSTGTGDPDPDTARKPKYE
IONQTLPVDFEALHLYRGLGDSRYTFCNGEKTIDKLOELGAEHFTDYGADVCV
GELAVEPWIAGLWPAIRKGFRRSNGQEBISGALPVASASRTDLVKSLELHIESQV
ELIARFDSGRKDSVTLKONAVNSQNSNVIEDFESSLTRSVPLSQASINIGLPEBY
LOVHIOESIGOESEOVSVTSADPVPOVPSIKAVOLTNDALKTLLVLEIDISTDPSY
OPGDAFSVYICPNSSDEVOILORLOLEDRHRCVLLKTKADTRKKGATLPHIPAGCS
LQPIFTWCEIRATIRKKAFLRLVDYTSASAKRRLQELCSQGAADYRPFYRDACAC
LIDLILAPPSQCPPLSLLEHLPKQAPRYSCASSLFPFGKLFHFNIVETPLATIT
EVLARKGVCQGMALIVASVLOPNIHASHDSKALAPKISIPRTTNSFHLDDPSIP
IIMVGPETIAPFIFLOHREKLOHDPGNFGAMWLPFGCHKRDYLFRRSLRHLF
KHGILTHLKVSFSRPAVGESEBAPAKYVDNIOHQGOVARILLQENGIYYCGDAKN
NAKDVDLALVQIISKEVGEKLEAKTILAEKERYLDIWS"
transit_peptide
31..123
/gene="MTRR"
/evidence=not_experimental
2209..3291
/gene="MTRR"
/evidence=not_experimental
3246..3251
/gene="MTRR"
/evidence=not_experimental
3258
/gene="MTRR"
/evidence=not_experimental
3276
/gene="MTRR"
/evidence=experimental
polyA_site
polyA_site
polyA_site
ORIGIN

```

Query Match 97.6%; Score 2046; DB 9; Length 3291;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2096; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGAGAGTTTCTGTACTATATGCTACACAGAGGAGCAGGCAAAAGCCATCGAGAA 60  
DB 112 ATAGAGAGTTTCTGTACTATATGCTACACAGAGGAGCAGGCAAAAGCCATCGAGAA 171  
QY 61 GAAATGTGTAGCAAGCTGTGTATCATGATTTTCTGAGATCTTCACTATATTAGTAA 120  
DB 172 GAAATGTGTAGCAAGCTGTGTATCATGATTTTCTGAGATCTTCACTATATTAGTAA 231  
QY 121 TCCGATTAAGTATGACCTTAACCAAGCAAGCTCTGTGTGTGTGTGTCTTACCAAG 180  
DB 232 TCCGATTAAGTATGACCTTAACCAAGCAAGCTCTGTGTGTGTGTGTCTTACCAAG 291  
QY 181 GGCACCGGAGACCCACCGACAGAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACA 240  
DB 292 GGCACCGGAGACCCACCGACAGAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACA 351  
QY 241 CTGCGCGGTGATTTCTTGTCTCACCTGCGGTATGGTTACTGGGTCTCGGTATTCAGAA 300  
DB 352 CTGCGCGGTGATTTCTTGTCTCACCTGCGGTATGGTTACTGGGTCTCGGTATTCAGAA 411  
QY 301 TACACCTACTTTTGCAATGGGGGGAAATATGATTAAGCACTTCAAGGCTTGGAGCC 360  
DB 412 TACACCTACTTTTGCAATGGGGGGAAATATGATTAAGCACTTCAAGGCTTGGAGCC 471  
QY 361 CGGCAATTTCTATGACACTGTGACATGACATGATGATGATGATTTAGAACTTGTGTGAG 420  
DB 472 CGGCAATTTCTATGACACTGTGACATGACATGATGATGATGATTTAGAACTTGTGTGAG 531  
QY 421 CCGTGATTTCTGACTCTGAGCACTGAGCAAGCTTCAAGAAAGATTTTATAGTCAAGAGACAA 480  
DB 532 CCGTGATTTCTGACTCTGAGCACTGAGCAAGCTTCAAGAAAGATTTTATAGTCAAGAGACAA 591  
QY 481 GAGGAGTAAGTGGCGGCACTCCGGTGGCATCACTGCACTCTGTGAGGAGAGACCTTGTG 540  
DB 592 GAGGAGTAAGTGGCGGCACTCCGGTGGCATCACTGCACTCTGTGAGGAGAGACCTTGTG 651  
QY 541 AAGTCAGAGCTGTACATCATTTGAATCTCAAGTCAGGCTTCTGAGATTCATGATTCAGGA 600  
DB 652 AAGTCAGAGCTGTACATCATTTGAATCTCAAGTCAGGCTTCTGAGATTCATGATTCAGGA 711  
QY 601 AGAAAGATTTCTGAGTTTGAAGCAAAATGACGTGAACAGCAACCAATCCAATGTGTGA 660  
DB 712 AGAAAGATTTCTGAGTTTGAAGCAAAATGACGTGAACAGCAACCAATCCAATGTGTGA 771  
QY 661 ATTGAAGCTTTGAGTCTGACTTAACCGGTTCGGTACCCGCACTCTCAAGAGCTCTGTG 720  
DB 772 ATTGAAGCTTTGAGTCTGACTTAACCGGTTCGGTACCCGCACTCTCAAGAGCTCTGTG 831  
QY 721 AATATTTCTGTGTTTACCCCGAGATATTTACAGGTACATCTGAGAGAGTCTCTTGGCCAG 780  
DB 832 AATATTTCTGTGTTTACCCCGAGATATTTACAGGTACATCTGAGAGAGTCTCTTGGCCAG 891  
QY 781 GAGGAAAGCCAGATATCTGTGACTTCAGAGATTCAGTATTTTCAAGTGCATTTCAAG 840  
DB 892 GAGGAAAGCCAGATATCTGTGACTTCAGAGATTCAGTATTTTCAAGTGCATTTCAAG 951  
QY 841 GCGATTCACTTACTAGATGATGCAATTAACCACTCTGCTGTGATGAATTTGGACATT 900  
DB 952 GCGATTCACTTACTAGATGATGCAATTAACCACTCTGCTGTGATGAATTTGGACATT 1011  
QY 901 TCAATATACAGACTTTTCTATCAGGCTGAGATGCTTCAAGGATCTGCTCCCTAACAT 960  
DB 1012 TCAATATACAGACTTTTCTATCAGGCTGAGATGCTTCAAGGATCTGCTCCCTAACAT 1071  
QY 961 GATTTGAGGTACAAAGCTTACTCCAAAGCTGCAAGTTGAAGATTAAGAGAGACTGC 1020  
DB 1072 GATTTGAGGTACAAAGCTTACTCCAAAGCTGCAAGTTGAAGATTAAGAGAGACTGC 1131  
QY 1021 GTGCTTTTGAATAAAGGAGACAAAGAAAGAGAGCTACTTACCCGACATATA 1080

DB 1132 GTGCTTTTGAATAAAGGAGACAAAGAAAGAGCTACTTACCCGACATATA 1191  
QY 1081 CCGCGGAGATGTTCTCTCCAGTTCAATTTTACCTGTGTCTTGAATTCGAGCAATTCCT 1140  
DB 1192 CCGCGGAGATGTTCTCTCCAGTTCAATTTTACCTGTGTCTTGAATTCGAGCAATTCCT 1251  
QY 1141 AAAAGGCAATTTTGGAGCCCTTGTGACATATACAGATGACAGTGTGAAAAAGCCAGG 1200  
DB 1252 AAAAGGCAATTTTGGAGCCCTTGTGACATATACAGATGACAGTGTGAAAAAGCCAGG 1311  
QY 1201 CTACAGAGCTGTGACATTAACAGAGGAGCCGATTAATAGCCGTTTGTAGAGATGCC 1260  
DB 1312 CTACAGAGCTGTGACATTAACAGAGGAGCCGATTAATAGCCGTTTGTAGAGATGCC 1371  
QY 1261 TGTGCTGTGTGTGATCTCTCTCGGTTTCCCTTCTTCCAGCCACCACTCACTCTC 1320  
DB 1372 TGTGCTGTGTGTGATCTCTCTCGGTTTCCCTTCTTCCAGCCACCACTCACTCTC 1431  
QY 1321 CTGCTGAAATCTTCTTAACTTCAACCAAGCAATTTCTGTGCAAGCTCAAGTTA 1380  
DB 1432 CTGCTGAAATCTTCTTAACTTCAACCAAGCAATTTCTGTGCAAGCTCAAGTTA 1491  
QY 1381 TTTCACCGAGAAAGCTCATTTTGTCTTCAACATTTGTGAATTTCTGTACTGCAACA 1440  
DB 1492 TTTCACCGAGAAAGCTCATTTTGTCTTCAACATTTGTGAATTTCTGTACTGCAACA 1551  
QY 1441 ACAGAGTTCTGCGGAAAGGAGATATGATGATGAGCTGTGCTGTGTGTGTCTTCAATT 1500  
DB 1552 ACAGAGTTCTGCGGAAAGGAGATATGATGATGAGCTGTGCTGTGTGTGTCTTCAATT 1561  
QY 1501 CTTCAGCCAAATCATGATGATCCATGAAAGACGCGGAAAGCCCTGTGCTTAAATTA 1560  
DB 1612 CTTCAGCCAAATCATGATGATCCATGAAAGACGCGGAAAGCCCTGTGCTTAAATTA 1671  
QY 1561 TCCATCTCTCTGAAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCCTATC 1620  
DB 1672 TCCATCTCTCTGAAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCCTATC 1731  
QY 1621 ATATAGTGGGTTCAGAAACCGGCAATAGCCCGTTTATTTGGTTCTTACAACTAGAGAG 1680  
DB 1732 ATATAGTGGGTTCAGAAACCGGCAATAGCCCGTTTATTTGGTTCTTACAACTAGAGAG 1791  
QY 1681 AAATCTCAAGAACACACCAAGATGGAATTTTGGACATGTGTGTGTGTGTGTGTGCTGC 1740  
DB 1792 AAATCTCAAGAACACACCAAGATGGAATTTTGGACATGTGTGTGTGTGTGTGTGCTGC 1851  
QY 1741 AGGCAATAGGATAGGATATATCTATTCAGAAAGAGCTCAGACATTTCTTAAGCATGGG 1800  
DB 1852 AGGCAATAGGATAGGATATATCTATTCAGAAAGAGCTCAGACATTTCTTAAGCATGGG 1911  
QY 1801 ATCTTAATCATCTAAAGGTTTCTTCTCAAGAGATGCTCTGTGGGAGAGAGAGCC 1860  
DB 1912 ATCTTAATCATCTAAAGGTTTCTTCTCAAGAGATGCTCTGTGGGAGAGAGAGCC 1971  
QY 1861 CCAAGCAAGTATGATCAAGACAACTCAGCTTCAATGCGCAGAGGTGGCAGAAATCTCTC 1920  
DB 1972 CCAAGCAAGTATGATCAAGACAACTCAGCTTCAATGCGCAGAGGTGGCAGAAATCTCTC 2031  
QY 1921 CTCCAGAGAAAGGCAATTTTATGTGTGTGAGATGCAAGAAATTTGGCAGAGATGTA 1980  
DB 2032 CTCCAGAGAAAGGCAATTTTATGTGTGTGAGATGCAAGAAATTTGGCAGAGATGTA 2091  
QY 1981 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGTTGAGTTGAAAACTAGAGCAATG 2040  
DB 2092 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGTTGAGTTGAAAACTAGAGCAATG 2151  
QY 2041 AAAACCTTGCCCATTTTAAAGAGAAAAAGCTACTTCAAGATATTTGTGTCATTA 2097  
DB 2152 AAAACCTTGCCCATTTTAAAGAGAAAAAGCTACTTCAAGATATTTGTGTCATTA 2208

RESULT 5

CQ726091 LOCUS 3241 bp DNA linear PAT 03-FEB-2004  
 DEFINITION Sequence 12025 from Patent WO02068579.  
 ACCESSION CQ726091  
 VERSION CQ726091.1 GI:42288134  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1  
 AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
 TITLE Kite, such as nucleic acid arrays, comprising a majority of  
 humanexons or transcripts, for detecting expression and other uses  
 thereof  
 JOURNAL Patent: WO 02068579-A 12025 06-SEP-2002;  
 PB Corporation  
 FEATURES  
 source 1..3241  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 ORIGIN  
 Query Match 92.7%; Score 1944; DB 6; Length 3241;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 2094; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 1 ATGAGAGAGGTTCTGTACTATATGCTACACAGCAGGACAGGCAAGGCCATCGAGAA 60  
 80 ATGAGAGAGGTTCTGTACTATATGCTACACAGCAGGACAGGCAAGGCCATCGAGAA 139  
 61 GAAATGTGAGCAAGCTGTGTGATGATGATTTTTCGAGATCTTCACTATATTTAGAA 120  
 140 GAAATGTGAGCAAGCTGTGTGATGATGATTTTTCGAGATCTTCACTATATTTAGAA 199  
 121 TCCGATTAAGTATGACCTTAACAAACGAAAGAGCTCTCTTGTGTTGTTGTTCTACAG 180  
 200 TCCGATTAAGTATGACCTTAACAAACGAAAGAGCTCTCTTGTGTTGTTGTTCTACAG 259  
 181 GGCACCGAGAGCCACCCGACACAGCCGCAAGTTGTTAAGAAATACAGAACCAACA 240  
 260 GGCACCGAGAGCCACCCGACACAGCCGCAAGTTGTTAAGAAATACAGAACCAACA 319  
 241 CTGCGGTTGATTTCTTCTGCTCACTGCGGTATGAGTTTCTGCTGCTGCTGCTGCTG 300  
 320 CTGCGGTTGATTTCTTCTGCTCACTGCGGTATGAGTTTCTGCTGCTGCTGCTGCTG 379  
 301 TACACCTACTTTTGCAATGAGGAGGAGATTAATTGATTAACGACTTCAAGAGCTTGAGCC 360  
 380 TACACCTACTTTTGCAATGAGGAGGAGATTAATTGATTAACGACTTCAAGAGCTTGAGCC 439  
 361 CGGCAATTTCTATGACACTGAGCAATGAGATGACTGTGATGATTTAGAACTTGTGTTGAG 420  
 440 CGGCAATTTCTATGACACTGAGCAATGAGATGACTGTGATGATTTAGAACTTGTGTTGAG 499  
 421 CCGTGAATGCTGAGACTCTGCGCAGCCCTCAAGAAAGATTTTATGTTAGCAGAGAGCA 480  
 500 CCGTGAATGCTGAGACTCTGCGCAGCCCTCAAGAAAGATTTTATGTTAGCAGAGAGCA 559  
 481 GAGAGATTAAGTATGAGCACTCCCGGTGAGATCACTGATCTCTTGAGAGCAGACTTGTG 540  
 560 GAGAGATTAAGTATGAGCACTCCCGGTGAGATCACTGATCTCTTGAGAGCAGACTTGTG 619  
 541 AAGTCAAGCTGCTACATTAATCTCAAGTCAAGCTTCTGAGATTTGATGATTTGAGAG 600  
 620 AAGTCAAGCTGCTACATTAATCTCAAGTCAAGCTTCTGAGATTTGATGATTTGAGAG 679  
 601 AAGAAAGATTTCTGAGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATTCATTTGTA 660  
 680 AAGAAAGATTTCTGAGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATTCATTTGTA 739  
 661 ATTGAAGATTTGAGTCTCATTAACCGGTGAGTACCCCACTCTCAAGAGCTCTG 720

Db 740 ATTGAAGATTTGAGTCTCATTAACCGGTGAGTACCCCACTCTCAAGAGCTCTG 799  
 721 AATATTTCTGTTTACCCCAAGATTTTACAGTATCACTGCAAGAGCTCTGTTCCAG 780  
 800 AATATTTCTGTTTACCCCAAGATTTTACAGTATCACTGCAAGAGCTCTGTTCCAG 859  
 781 GAGAAAGCCAAATTTCTGAGCTTCAAGATTCAGAGATCCAGTTTCAAGTCAATTTCAAG 840  
 860 GAGAAAGCCAAATTTCTGAGCTTCAAGATTCAGAGATCCAGTTTCAAGTCAATTTCAAG 919  
 841 GAGATTCAACTTACTAGCAATGATGATTAACCACTCTGCTGATGATTTGAGAGCTT 900  
 920 GAGATTCAACTTACTAGCAATGATGATTAACCACTCTGCTGATGATTTGAGAGCTT 979  
 901 TCAATTAAGCACTTTTCTCATGAGCTGAGATGCTTCAAGGATGATTCAGCTTACAGT 960  
 980 TCAATTAAGCACTTTTCTCATGAGCTGAGATGCTTCAAGGATGATTCAGCTTACAGT 1039  
 961 GATTCAGATTAACCAAGCTTCAAGATTCAGAGCTTGAAGATTAAGAGAGCACTG 1020  
 1040 GATTCAGATTAACCAAGCTTCAAGATTCAGAGCTTGAAGATTAAGAGAGCACTG 1099  
 1021 GTCTTTTGAATAAAGGAGAGCAACAAAGAAAGAGCTTACCTTACCCAGCATATA 1080  
 1100 GTCTTTTGAATAAAGGAGAGCAACAAAGAAAGAGCTTACCTTACCCAGCATATA 1159  
 1081 CCGTGGAGATGTTTCTCTCAAGTTCATTTTACCTGCTGCTTGAAGATTCAGAGCAATTC 1140  
 1160 CCGTGGAGATGTTTCTCTCAAGTTCATTTTACCTGCTGCTTGAAGATTCAGAGCAATTC 1219  
 1141 AAAAAAGCATTTTGTGAGAGCCCTGCTGAGCTATACAGAGAGCTGTAAGAAAGCAGAG 1200  
 1220 AAAAAAGCATTTTGTGAGAGCCCTGCTGAGCTATACAGAGAGCTGTAAGAAAGCAGAG 1279  
 1201 CTACAGAGCTGTGAGTAAACAAGGAGGAGCCGATTAATAGCCGCTTTTACAGAGATGCC 1260  
 1280 CTACAGAGCTGTGAGTAAACAAGGAGGAGCCGATTAATAGCCGCTTTTACAGAGATGCC 1339  
 1261 TGTGCTGCTGTTGATGATCTCTCTGCTTCCCTTCTTCCAGAGCAGCAGCTCAAGTCTC 1320  
 1340 TGTGCTGCTGTTGATGATCTCTCTGCTTCCCTTCTTCCAGAGCAGCAGCTCAAGTCTC 1399  
 1321 CTGCTGAGCATCTTCTTAACTTGAACCAAGCAATTCGATGAGTCTCAAGTCTTAA 1380  
 1400 CTGCTGAGCATCTTCTTAACTTGAACCAAGCAATTCGATGAGTCTCAAGTCTTAA 1459  
 1381 TTTCAACCAAGAAAGCTCAATTTGCTTCAATTTGAGAAATTTCTGCTCACTGAGCA 1440  
 1460 TTTCAACCAAGAAAGCTCAATTTGCTTCAATTTGAGAAATTTCTGCTCACTGAGCA 1519  
 1441 ACGAGGTTCTCGGAGGAGATGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
 1520 ACGAGGTTCTCGGAGGAGATGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1579  
 1501 CTTCAGCCAAATATCATGATCCCATGAAAGCAGGAGGAAAGCCCTGCTCTTAAGATA 1560  
 1580 CTTCAGCCAAATATCATGATCCCATGAAAGCAGGAGGAAAGCCCTGCTCTTAAGATA 1639  
 1561 TCCATCTCTCTGAGCAACAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
 1640 TCCATCTCTCTGAGCAACAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1699  
 1621 ATTAATGATGATGATGAGAAACCGGATAGCCCGGTTATTTAGGTTCTTCAACAATGAGAG 1759  
 1700 ATTAATGATGATGATGAGAAACCGGATAGCCCGGTTATTTAGGTTCTTCAACAATGAGAG 1759  
 1681 AATCTCAAGAAACAACCCAGATGAAATTTTGAAGCAATGATGATTTTGTGCTGCTC 1740  
 1760 AATCTCAAGAAACAACCCAGATGAAATTTTGAAGCAATGATGATTTTGTGCTGCTC 1819  
 1741 AAGCATTAAGATGAGGATTAATCTATTCAGAAAGAGCTCAGACATTTCTTAAAGATGAG 1800



Db 1820 AGCATRAGGATGGGATTTATCTATTCAGAAAAAGCTCAGACATTTCTTAAAGCATGG 1879

Qy 1801 ATCTTAATCATCTTAAAGGTTTCCTTCTCAAGAGATGCTCTGTTGGGAGAGAAAGCC 1860

Db 1880 ATCTTAATCATCTTAAAGGTTTCCTTCTCAAGAGATGCTCTGTTGGGAGAGAAAGCC 1939

Qy 1861 CCGCAAAAGTATGTACAGACACATCTCAGCTTCATGCGACGAGTGGCGAAGATCTTC 1920

Db 1940 CCGCAAAAGTATGTACAGACACATCTCAGCTTCATGCGACGAGTGGCGAAGATCTTC 1999

Qy 1921 CTCAGAGAGACGGCCATATTTATGTTGTGAGATGCAAGAAATATGCGCAAGATGTA 1980

Db 2000 CTCAGAGAGACGGCCATATTTATGTTGTGAGATGCAAGAAATATGCGCAAGATGTA 2059

Qy 1981 CATGATGCCCTTGTGCAATATATAGCAAGAGTTGAGTTGAAAACTAGAGCAATG 2040

Db 2060 CATGATGCCCTTGTGCAATATATAGCAAGAGTTGAGTTGAAAACTAGAGCAATG 2119

Qy 2041 AAAACCTGGCCACTTTAAAAAGAAAAAGCGTACCTTCAGATATTTGTCTATTA 2097

Db 2120 AAAACCTGGCCACTTTAAAAAGAAAAAGCGTACCTTCAGATATTTGTCTATTA 2176

RESULT 6  
BC054816  
LOCUS  
DEFINITION  
Homo sapiens 5-methyltetrahydrofolate-homocysteine  
methyltransferase reductase, mRNA (cDNA clone IMAGE:5205285),  
partial cds.  
BC054816  
VERSION  
BC054816.1 GI:33392775  
KEYWORDS  
Homo sapiens (human)  
SOURCE  
Homo sapiens  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS  
1 (bases 1 to 3310)  
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,  
Klausner,R.D., Collins,F.S., Wagner,J., Shenmen,C.M., Schuler,G.D.,  
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,  
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,  
Diatchenko,L., Marisina,K., Farmer,A.A., Rubin,G.M., Hong,L.,  
Sachetron,M., Soares,W.B., Bonaldo,M.F., Casavant,I.L.,  
Schaefer,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,  
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,  
Abramsen,R.D., Mullany,S.J., Bosak,S.A., McSwan,P.J.,  
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,  
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,  
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,  
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,  
Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,  
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,  
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,  
Butterfield,Y.S., Krzywicki,M.T., Skalska,U., Smailus,D.E.,  
Schmeich,A., Schein,J.B., Jones,S.V. and Marra,M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
22388257  
12477932  
2 (bases 1 to 3310)  
Strausberg,R.  
Direct Submission  
Submitted (03-JUL-2003) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [gcgbs-remail.nih.gov](mailto:gcgbs-remail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)

REMARK  
COMMENT

DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc\\_mgc@hgrl.nih.gov](mailto:nisc_mgc@hgrl.nih.gov)

Akhter,N., Ayale,K., Beckstrom-Sternberg,S.M., Benjamin,B.,  
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,  
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,  
Hansen,N., Ho,S.-L., Karlina,E., Kwong,P., Laric,P., Legaapi,R.,  
Maduro,O.L., Mastello,C., Maskeri,B., Mastrian,S.D., McLooney,J.C.,  
McDowell,J., Pearson,R., Stantipop,S., Thomas,P.J., Touchman,J.W.,  
Toungian,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,  
Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LNLN at: <http://image.lnl.gov>  
Series: IRAC Plate: 115 Row: d Column: 11  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 4505278.

FEATURES  
source  
1. .3310  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5205285"  
/tissue\_type="lung, spleen, fetal, pooled"  
/clone\_lib="NIH MGC\_122"  
/lab\_host="DH10B"  
/note="Vector: pCMV-Sport6"  
1. .3310  
/gene="MTRR"  
/note="synonym: MSR"  
/db\_xref="LOCUSID:4552"  
/db\_xref="MIM:602568"  
1. .2149  
/gene="MTRR"  
/codon\_start=2  
/product="MTRR protein"  
/protein\_id="AAH54816.1"  
/db\_xref="GI:33392776"  
/db\_xref="LOCUSID:4552"  
/db\_xref="MIM:602568"  
/translation="VARLARGFTTCTLEWRRPFLILRTGCGAKATAEICEQAV  
VHGFSLDHCISSDYDKLTETAPLVAVSTGTDPDTARKEVKEIQNQLPVD  
FALRYGLGLGDSSEYTCNGEKIDKLOELGAGHFDTGADDCVLELVBEPI  
AGLMPALRGFRSSRCOEISGALPVASPASSTPDVKSLELHESDVELLRDSDGR  
KQSEVYKQNAVNSNOSNVITDEPSLTSTSVPLSOASINIPGLPREYOVHQSIG  
QBSQSVTSADPVPOVPYPSKAVCLTNDATKTTLLVLEIDSYSDSGDPAFSYC  
PNSDSVQSLQRLQEDREKHCVLKIKADTKKATLPQHPACSLQFITTWCLB  
IRAIIPKAEFLRALVDYTSASERKRLQELCSKQADYSPVADACGLDLILAPS  
CCPPLSLLEHLPKQPRPYSCASSLFFHGKLFVNFVLESTATTETLRGVCTG  
WIALIVASYLQPNHSHSDGSLAKLPKISISPTNSFVLPDPPIPIIMVPGTGI  
APRIGLQREKIOEHOHPDNGFAMMLPGCRHKDDYERKELRHLKHGILTHLYKV  
SFRDPAVGEERPAKYVDNITDHCQVARIILQNGHIIYVGDAKNAKVDHALV  
QIISKVGEVKEAMKTLATLKEKRTLDVINS"  
59. .493  
/gene="MTRR"  
/note="flavodoxin; Region: Flavodoxin"  
/db\_xref="CD:pfam0256"  
848. .2143  
/gene="MTRR"  
/note="CysJ; Region: Sulfite reductase, alpha subunit  
(flavo-protein) (Inorganic ion transport and metabolism)"  
/db\_xref="CD:COG0369"

ORIGIN  
Query Match 90.3%; Score 1893; DB 9; Length 3310;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2093; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 ATAGAGAGGTTTCTGTTACTATATGCTACACACAGGAGAGCAAGGCATCGCAGAA 60  
Db 53 ATAGAGAGGTTTCTGTTACTATATGCTACACACAGGAGAGCAAGGCATCGCAGAA 112



61 GAATGTGTGAGCAAGCTGTGGTACATGATTTTCTGAGATCTTCAATATATAGTGA 120  
113 GAAATATGTGAGCAAGCTGTGGTACATGATTTTCTGAGATCTTCAATATATAGTGA 172  
121 TCCGATTAAGTATGACCTTAATAAACCAGAAAGCTCTCTTGTGTGTGTGTCTACACG 180  
173 TCCGATTAAGTATGACCTTAATAAACCAGAAAGCTCTCTTGTGTGTGTGTCTACACG 232  
181 GGCACCGGAGACCCACCCGACACAGCCCGCAAGTTGTTAAGAAATACAAACCAACA 240  
233 GGCACCGGAGACCCACCCGACACAGCCCGCAAGTTGTTAAGAAATACAAACCAACA 292  
241 CTGCGGTGATTTCTTGTGTGACCTGCGGTATGGGTATCTGGGTCTCGGTATTCAGA 300  
293 CTGCGGTGATTTCTTGTGTGACCTGCGGTATGGGTATCTGGGTCTCGGTATTCAGA 352  
301 TACACCTACTTTTGTGAAGTGGGGGAGATTAATGATTAACAAGCTTCAGAGCTTGAGCC 360  
353 TACACCTACTTTTGTGAAGTGGGGGAGATTAATGATTAACAAGCTTCAGAGCTTGAGCC 412  
361 CCGCATTTCTATGACACTGGAACATGCAATGACTGTGTAGTTTAAACTTGTGTGAG 420  
413 CCGCATTTCTATGACACTGGAACATGCAATGACTGTGTAGTTTAAACTTGTGTGAG 472  
421 CCGTGAATGTGAGACTCTGCGGACGCTCAGAAAGCATTTTATGTCAGAGAGAGCAA 480  
473 CCGTGAATGTGAGACTCTGCGGACGCTCAGAAAGCATTTTATGTCAGAGAGAGCAA 532  
481 GAGAGATATAGTGGCGCACTCCCGGTGGATCACTGCACTCTTGAAGACAGACTTGTG 540  
533 GAGAGATATAGTGGCGCACTCCCGGTGGATCACTGCACTCTTGAAGACAGACTTGTG 592  
541 AAGTCAGAGCTGCTACATTTGATTCATGATCTGAGCTTCTGATTCATGATTCAGAA 600  
593 AAGTCAGAGCTGCTACATTTGATTCATGATCTGAGCTTCTGATTCATGATTCAGAA 652  
601 AGAAGAGATTTGAGGTTTGAAGCAAAATGACAGTAAACAGCAACCAATCCATGTTGTA 660  
653 AGAAGAGATTTGAGGTTTGAAGCAAAATGACAGTAAACAGCAACCAATCCATGTTGTA 712  
661 ATTGAAGACTTTGAGTCTCACTTACCCGTTCCGTTACCCCACTCTACAGAGCTCTCTG 720  
713 ATTGAAGACTTTGAGTCTCACTTACCCGTTCCGTTACCCCACTCTACAGAGCTCTCTG 772  
721 AATATTCCTGTGTACCCCAAGAAATTTACAGATACATCTGAGAGAGTCTCTTGCCAG 780  
773 AATATTCCTGTGTACCCCAAGAAATTTACAGATACATCTGAGAGAGTCTCTTGCCAG 832  
781 GAGGAAAGCCAGATATCTGTGACTTACAGAGATCCAGTTTTCATGATCCATTTCAAG 840  
833 GAGGAAAGCCAGATATCTGTGACTTACAGAGATCCAGTTTTCATGATCCATTTCAAG 892  
841 GAGGAAAGCCAGATATCTGTGACTTACAGAGATCCAGTTTTCATGATCCATTTCAAG 900  
893 GAGGAAAGCCAGATATCTGTGACTTACAGAGATCCAGTTTTCATGATCCATTTCAAG 952  
901 TCAATATACAGACTTTTCTATACAGCTGAGATGCTTCAAGCTGTATCTGCTTAAACAGT 960  
953 TCAATATACAGACTTTTCTATACAGCTGAGATGCTTCAAGCTGTATCTGCTTAAACAGT 1012  
961 GATTCGAGGTACAAAGCTCTTCAAGAGCTGAGATGCTTCAAGATTAAGAAAGAGCACTGC 1020  
1013 GATTCGAGGTACAAAGCTCTTCAAGAGCTGAGATGCTTCAAGATTAAGAAAGAGCACTGC 1072  
1021 GTCCTTTTGAATTAAGAGAGACACAAAGAAAGAGAGCTACTTACCCAGCATATA 1080  
1073 GTCCTTTTGAATTAAGAGAGACACAAAGAAAGAGAGCTACTTACCCAGCATATA 1132  
1081 CTGTGCGGAGATGTTCTCTCAAGTTCAATTTTACCTGTGTCTTGAATTCGAGCAATTTCT 1140  
1133 CTGTGCGGAGATGTTCTCTCAAGTTCAATTTTACCTGTGTCTTGAATTCGAGCAATTTCT 1192

1141 AAAAGGCAATTTTTCGAGCCCTTGTGACATATACAGTGTGAGAGAGAGAGAGAG 1200  
1193 AAAAGGCAATTTTTCGAGCCCTTGTGACATATACAGTGTGAGAGAGAGAGAGAG 1252  
1201 CTACAGAGCTGTGACATTAACAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260  
1253 CTACAGAGCTGTGACATTAACAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1312  
1261 TGTGCTGTGTGTGATCTCTCTGCTTCCCTTCTGTCAGAGAGAGAGAGAGAGAG 1320  
1313 TGTGCTGTGTGTGATCTCTCTGCTTCCCTTCTGTCAGAGAGAGAGAGAGAGAG 1372  
1321 CTGCTGAAACATCTTCTTAACCTTCAACCCAGAGAGAGAGAGAGAGAGAGAGAG 1380  
1373 CTGCTGAAACATCTTCTTAACCTTCAACCCAGAGAGAGAGAGAGAGAGAGAGAG 1432-  
1381 TTTTACCCAG 1440  
1433 TTTTACCCAG 1492  
1441 AGAGAGGTTCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500  
1493 AGAGAGGTTCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1552  
1501 CTTCAGCCAAACATATCATCATCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560  
1553 CTTCAGCCAAACATATCATCATCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1612  
1561 TCCATCTCTCTGAG 1620  
1613 TCCATCTCTCTGAG 1672  
1621 ATTAAGTGTGTGTCAG 1680  
1673 ATTAAGTGTGTGTCAG 1732  
1681 AAATCTCAAG 1740  
1733 AAATCTCAAG 1792  
1741 AGGATTAAG 1800  
1793 AGGATTAAG 1852  
1801 ATCTTAATCTATCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860  
1853 ATCTTAATCTATCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1912  
1861 CCAGCAAGATATGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920  
1913 CCAGCAAGATATGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1972  
1921 CTTCAG 1980  
1973 CTTCAG 2032  
1981 CATGATCCCTTGTGAG 2040  
2033 CATGATCCCTTGTGAG 2092  
2041 AAAAGCCCTGAG 2097  
2093 AAAAGCCCTGAG 2149

RESULT 7  
BV177620/c 2933 bp DNA linear STS 10-JUN-2004  
LOCUS BV177620 Human DNA (Sequenced) Homo sapiens STS genomic, sequence  
DEFINITION sqm85800 Human DNA (Sequenced) Homo sapiens STS genomic, sequence  
tagged site.  
ACCESSION BV177620  
VERSION BV177620.1 GI:48013757  
KEYWORDS STS.

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 2933)  
AUTHORS Nelson,R.M., Marnellos,G., Kemmerer,S., Hoyal,C.R., Shi,M.M.,  
Cantor,C.R. and Braun,A.  
TITLE Large-Scale Validation of Single Nucleotide Polymorphisms in Gene  
Regions  
JOURNAL Genome Res. (2004) In press  
COMMENT Contact: Andreas Braun  
Pharmaceuticals division  
Sequenom, Inc.  
3595 John Hopkins Court, San Diego, CA 92121, USA  
Tel: 18582029018  
Fax: 18582029020  
Email: abraun@sequenom.com  
Primer A: No primer sequence submitted  
Primer B: No primer sequence submitted  
STS size: 2933.

FEATURES  
source Location/Qualifiers  
1..2933  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone\_lib="Human DNA (Sequenom)"  
<1..>2933

ORIGIN  
STS

Query Match 63.8%; Score 1338; DB 11; Length 2933;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1778; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY 315 CAATGGGGGAGATTAATTAACGACTTCAAGACTTGAGCCCGCATTTCTATGA 374  
Db 2846 CAATGGGGGAGATTAATTAACGACTTCAAGACTTGAGCCCGCATTTCTATGA 2787  
QY 375 CACTGCAGATGACAGATGCTGTAGGTTTAACTTGTGGTTGAGCCGTGATTCCTG 434  
Db 2786 CACTGCAGATGACAGATGCTGTAGGTTTAACTTGTGGTTGAGCCGTGATTCCTG 2727  
QY 435 ACCTGGCCAGCCCTCGAAGAAAGCATTTTAGTCAAGAGAGACAAGAGAGATTAATGG 494  
Db 2726 ACCTGGCCAGCCCTCGAAGAAAGCATTTTAGTCAAGAGAGACAAGAGAGATTAATGG 2667  
QY 495 CGCACTCCCGGTGACATCACTGCATCTTGAAGACAGACCTTGTGAAGTCAGAGCTGCT 554  
Db 2666 CGCACTCCCGGTGACATCACTGCATCTTGAAGACAGACCTTGTGAAGTCAGAGCTGCT 2607  
QY 555 ACAATTTGAATCTCAAGTGAAGCTTCTGAGATTGATTCAGAGAAAGAGATTCGA 614  
Db 2606 ACAATTTGAATCTCAAGTGAAGCTTCTGAGATTGATTCAGAGAAAGAGATTCGA 2547  
QY 615 GGTTTTGAAGCAAAATGCAATGAAACCAATCCAAATGTTTATTTGAAGCTTGA 674  
Db 2546 GGTTTTGAAGCAAAATGCAATGAAACCAATCCAAATGTTTATTTGAAGCTTGA 2487  
QY 675 GTCTCACTTAACCGTTCGATACCCCACTCTCAAGACCTCTCTGAATATTCCTGCTT 734  
Db 2486 GTCTCACTTAACCGTTCGATACCCCACTCTCTGAATATTCCTGCTT 2427  
QY 735 ACCCCAGAAATATTTACAGGTACATCTGACAGAGCTCTTGGCCAGAGAGAAACCAAGT 794  
Db 2426 ACCCCAGAAATATTTACAGGTACATCTGACAGAGCTCTTGGCCAGAGAGAAACCAAGT 2367  
QY 795 ATCTGTGACTTCAGACATCCAGTTTTCAGTGCCTCAATTTCAAGGACATTCATCTAC 854  
Db 2366 ATCTGTGACTTCAGACATCCAGTTTTCAGTGCCTCAATTTCAAGGACATTCATCTAC 2307  
QY 855 TACGAATGATGCGCATTAATAAACCATCTGCTGTGTAATGGAATTCATTAATACAGATT 914  
Db 2306 TACGAATGATGCGCATTAATAAACCATCTGCTGTGTAATGGAATTCATTAATACAGATT 2247

QY 915 TTCTATCAAGCTGAGATGCTTCAAGCTGATCTGCTTAACAGTATTCGAGGTACA 974  
Db 2246 TTCTATCAAGCTGAGATGCTTCAAGCTGATCTGCTTAACAGTATTCGAGGTACA 2187  
QY 975 AAGCCATCTCAAGCTGAGCTGATTAAGTAAAGAGAGACACTGCTCTTTGAAAT 1034  
Db 2186 AAGCCATCTCAAGCTGAGCTGATTAAGTAAAGAGAGACACTGCTCTTTGAAAT 2127  
QY 1035 AAGGACGACACAAAGAAAGAAAGAGACTTACCTCCAGCATATACCTGCGGATGTT 1094  
Db 2126 AAGGACGACACAAAGAAAGAAAGAGACTTACCTCCAGCATATACCTGCGGATGTT 2067  
QY 1095 TCTCCAGTTCAATTTTAACTGCTGCTTGAATCCAGCAATTCCTTAAAGGCAATTTT 1154  
Db 2066 TCTCCAGTTCAATTTTAACTGCTGCTTGAATCCAGCAATTCCTTAAAGGCAATTTT 2007  
QY 1155 GCGAGCCCTTGGAGATTAACAGTACAGTGTGTAAGGCGAGGCTACAGAGCTGTG 1214  
Db 2006 GCGAGCCCTTGGAGATTAACAGTGTGTAAGGCGAGGCTACAGAGCTGTG 1947  
QY 1215 CAGTAAACAAAGGGGACGCGATTAATAGCCGCTTGTACAGATGCTGTGCTTGT 1274  
Db 1946 CAGTAAACAAAGGGGACGCGATTAATAGCCGCTTGTGTACAGATGCTGTGCTTGT 1887  
QY 1275 GGATCTCTCTGCTGCTTTCCTTCTTCCAGCCACATCAATCTCTGTCGAACATCT 1334  
Db 1886 GGATCTCTCTGCTGCTTTCCTTCTTCCAGCCACATCAATCTCTGTCGAACATCT 1827  
QY 1335 TCCATAACTCAACCCAGACCATATTCGTGTGCAAGCTCAAGTTTATTCACCCAGAAA 1394  
Db 1826 TCCATAACTCAACCCAGACCATATTCGTGTGCAAGCTCAAGTTTATTCACCCAGAAA 1767  
QY 1395 GCTCCATTTTGTCTTCAACATGTGGAATTTCTGTACTCCACAAACAGAGTTCTG 1454  
Db 1766 GCTCCATTTTGTCTTCAACATGTGGAATTTCTGTACTCCACAAACAGAGTTCTG 1707  
QY 1455 GAAGGAGATGATGACAGGCTGCTGCTTGTGTTGTTGCTTCAAGCCAAACAT 1514  
Db 1706 GAAGGAGATGATGACAGGCTGCTGCTTGTGTTGTTGCTTCAAGCCAAACAT 1647  
QY 1515 ACATGATCCCATGAAGAGAGC - GGGAAAGCCCTGCTCTTAAGATATTCATCTCTC 1573  
Db 1646 ACATGATCCCATGAAGAGAGCAGCGGAAAGCCCTGCTCTTAAGATATTCATCTCTC 1587  
QY 1574 GAACAAACAATTTCTTCACTTACAGATGACCCCTCAATCCCATCATATGATGCTG 1633  
Db 1586 GAACAAACAATTTCTTCACTTACAGATGACCCCTCAATCCCATCATATGATGCTG 1527  
QY 1634 CAGGAACCGGATAGCCCGTATTTATGGTTCTTACACATTAAGAGAAATCTCAAGAAC 1693  
Db 1526 CAGGAACCGGATAG - CCGTATTTATGGTTCTTACACATTAAGAGAAATCTCAAGAAC 1468  
QY 1694 AACACCAATGGAATTTTGGAGCAATGATGTTTGGCTGAGGCAATTAAGATA 1753  
Db 1467 AACACCAATGGAATTTTGGAGCAATGATGTTTGGCTGAGGCAATTAAGATA 1408  
QY 1754 GGGATTATCTATCAGAAAAGAGCTGACATTTCTTAAAGCATGGAATCTTAATCATC 1813  
Db 1407 GGGATTATCTATCAGAAAAGAGCTGACATTTCTTAAAGCATGGAATCTTAATCATC 1348  
QY 1814 TAAAGTTTCTTCTCAAGATGCTCTGTTGGGAGAGAGAAAGCCCAAGCAATATG 1873  
Db 1347 TAAAGTTTCTTCTCAAGATGCTCTGTTGGGAGAGAGAAAGCCCAAGCAATATG 1288  
QY 1874 TACAAGCAACATCAGCTTCAATGACAGAGGTGAGAAATCTCTCCAGAGAAAG 1933  
Db 1287 TACAAGCAACATCAGCTTCAATGACAGAGGTGAGAAATCTCTCCAGAGAAAG 1228  
QY 1934 GCCATATTTATGCTGTGAGATGCAAGATATGACCAAGATGATCATGATGCTCTT 1993  
Db 1227 GCCATATTTATGCTGTGAGATGCAAGATATGACCAAGATGATCATGATGCTCTT 1168

QY 1994 TCACAATATATAGCAAGAGGTTGAGTTGAAAACTAGAAAGCAATGAAACCTGCGCA 2053  
DB 1167 TCACAATATATAGCAAGAGGTTGAGTTGAAAACTAGAAAGCAATGAAACCTGCGCA 1108  
QY 2054 CTTTAAAGAGAAAAAGCTACCTTACCTTACAGATATTTGTGATTA 2097  
DB 1107 CTTTAAAGAGAAAAAGCTACCTTACCTTACAGATATTTGTGATTA 1064

RESULT 8  
BV178010/c 2933 bp DNA linear STS 10-JUN-2004  
LOCUS sqm97986 Human DNA (Sequencem) Homo sapiens STS genomic, sequence  
DEFINITION tagged site.  
ACCESSION BV178010 GI:48014252  
VERSION BV178010.1  
KEYWORDS STS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 2933)  
Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,  
Cantor,C.R. and Braun,A.  
Large-Scale Validation of Single Nucleotide Polymorphisms in Gene  
Regions  
Genome Res. (2004) In press

JOURNAL COMMENT  
Contact: Andreas Braun  
Pharmaceuticals division  
Sequencem, Inc.  
3595 John Hopkins Court, San Diego, CA 92121, USA  
Tel: 18582029018  
Fax: 18582029020  
Email: abraun@sequencem.com  
Primer A: No primer sequence submitted  
Primer B: No primer sequence submitted  
STS size: 2933

FEATURES  
source Location/Qualifiers  
1..2933  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone\_jib="Human DNA (Sequencem)"  
<1..>2933

ORIGIN  
STS  
Query Match 63.8%; Score 1338; DB 11; Length 2933;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1778; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY 315 CAATGGGGGAGAGATATGATTAACGACTTCAAGAGCTTGGAGCCCGCATTTCTATGA 374  
DB 2846 CAATGGGGGAGAGATATGATTAACGACTTCAAGAGCTTGGAGCCCGCATTTCTATGA 2787  
QY 375 CACTGGAACATGACGATGACCTGCTAGCTTAAAGAACTTGGCTTGAACCCCTGGA 434  
DB 2786 CACTGGAACATGACGATGACCTGCTAGCTTAAAGAACTTGGCTTGAACCCCTGGA 2727  
QY 435 ACTCTGGCAGCCCTCAGAAAGATTTTAAAGTCAACAGAGACAGACAGAGATTAAGTGG 494  
DB 2726 ACTCTGGCAGCCCTCAGAAAGATTTTAAAGTCAACAGAGACAGACAGAGATTAAGTGG 2667  
QY 495 CGCAGCTCCGGTGGATCACTCTGATCTTGAAGACAGACCTTTGTAAGTCAAGCTGCT 554  
DB 2666 CGCAGCTCCGGTGGATCACTCTGATCTTGAAGACAGACCTTTGTAAGTCAAGCTGCT 2607  
QY 555 ACACATTTGATCTCAAGTGAAGCTTGAAGATTCAGATGATTCAGAAAGAAAGATTTCTGA 614  
DB 2606 ACACATTTGATCTCAAGTGAAGCTTGAAGATTCAGATGATTCAGAAAGAAAGATTTCTGA 2547  
QY 615 GGTTTTGAAGCAAAATGACAGTGAACAGCAACCAATGTTGTAATGAAAGACTTTGA 674

DB 2546 GGTTTTGAAGCAAAATGACAGTGAACAGCAACCAATGTTGTAATGAAAGACTTTGA 2487  
QY 675 GTTCCTCACTTACCCGTTGGTATCCCACTCTCAGAAAGCTCTGAATATTCCTGTTT 734  
DB 2486 GTTCCTCACTTACCCGTTGGTATCCCACTCTCAGAAAGCTCTGAATATTCCTGTTT 2427  
QY 735 ACCCCAGAAATATTTACAGATACATCTGAGAGAGTCTCTTGGCCAGAGAAAGCCAACT 794  
DB 2426 ACCCCAGAAATATTTACAGATACATCTGAGAGAGTCTCTTGGCCAGAGAAAGCCAACT 2367  
QY 795 ATCTGTACTTCAGACATCCAGTTTTCAGTCCCAATTTCAAAAGCCGTTCACTTAC 854  
DB 2366 ATCTGTACTTCAGACATCCAGTTTTCAGTCCCAATTTCAAAAGCCGTTCACTTAC 2307  
QY 855 TAGAATGATGCAATTAAGAAACACCTCTGCTGATGAATTTGACATTTCAATACAGACTT 914  
DB 2306 TAGAATGATGCAATTAAGAAACACCTCTGCTGATGAATTTGACATTTCAATACAGACTT 2247  
QY 915 TTCTATCAGCCTGAGAGATGCTTCAAGCTGATCTGCCCTTAACAGTATTTGAGGTACA 974  
DB 2246 TTCTATCAGCCTGAGAGATGCTTCAAGCTGATCTGCCCTTAACAGTATTTGAGGTACA 2187  
QY 975 AAGCTTACTCCAAAGACTGACGCTTGAAGATTAAGAGAGCACTGCTCTTTGAAAT 1034  
DB 2186 AAGCTTACTCCAAAGACTGACGCTTGAAGATTAAGAGAGCACTGCTCTTTGAAAT 2127  
QY 1035 AAGGAGACACAAAGAAAGAGAGTACCTTACCCGACATATACCTGCGGAGATGTTT 1094  
DB 2126 AAGGAGACACAAAGAAAGAGAGTACCTTACCCGACATATACCTGCGGAGATGTTT 2067  
QY 1095 TCTCAGTTCATTTTACCTGCTGCTTGAATTCGAGCAATTCCTAAAGGCAATTTT 1154  
DB 2066 TCTCAGTTCATTTTACCTGCTGCTTGAATTCGAGCAATTCCTAAAGGCAATTTT 2007  
QY 1155 GCGAGCCTTGTGACATATACAGTGAACGCTGTAAGAAAGCCAGGCTACAGAGCTGTG 1214  
DB 2006 GCGAGCCTTGTGACATATACAGTGAACGCTGTAAGAAAGCCAGGCTACAGAGCTGTG 1947  
QY 1215 CAGTAAACAAAGGGAGCCGATTAATAGCCGCTTTTGAAGAGTCCCTGCGCTGTT 1274  
DB 1946 CAGTAAACAAAGGGAGCCGATTAATAGCCGCTTTTGAAGAGTCCCTGCGCTGTT 1887  
QY 1275 GGATCTCTCTCTGCTTCCCTTCTTCCAGGACCACTCAGTCTCTCTGAAATCT 1334  
DB 1886 GGATCTCTCTCTGCTTCCCTTCTTCCAGGACCACTCAGTCTCTCTGAAATCT 1827  
QY 1335 TCTTAACTTCAACCCAGACATATTCGTGCAAGCTCAAGTTTATTCACCCAGAAA 1394  
DB 1826 TCTTAACTTCAACCCAGACATATTCGTGCAAGCTCAAGTTTATTCACCCAGAAA 1767  
QY 1395 GCTCCATTTTGTCTTCAACATTTGTGGAATTTCTGTCTACGCAACAGAGGTTCTGG 1454  
DB 1766 GCTCCATTTTGTCTTCAACATTTGTGGAATTTCTGTCTACGCAACAGAGGTTCTGG 1707  
QY 1455 GAAGGAGATGATACAGGCTGCTGCTTGTGTTGCTTCAAGTCTTCAAGCAAAAT 1514  
DB 1706 GAAGGAGATGATACAGGCTGCTGCTTGTGTTGCTTCAAGTCTTCAAGCAAAAT 1647  
QY 1515 ACATGATCCCATGAAGACAGC-GGGAAAGCCCTGCTCTTAAGATATCATCTCTCTC 1573  
DB 1646 ACATGATCCCATGAAGACAGCAGGAGAAAGCCCTGCTCTTAAGATATCATCTCTCTC 1587  
QY 1574 GAACAACAAAATTTTCCATTCACAGATGACCCCTCAATCCCATCAATATGATGAGTTC 1633  
DB 1586 GAACAACAAAATTTTCCATTCACAGATGACCCCTCAATCCCATCAATATGATGAGTTC 1527  
QY 1634 CAGAAACCGGATAGCCCGTTATTTGGTCTTCAACATATGAGAGAACTCAAGAAC 1693  
DB 1526 CAGAAACCGGATAGC-CGCTTTATTTGGTCTTCAACATATGAGAGAACTCAAGAAC 1468  
QY 1694 AACACCCAGATGAATTTTGAAGCAATGCTGTTTGTGCTGACGACATTAAGATA 1753  
DB 1467 AACACCCAGATGAATTTTGAAGCAATGCTGTTTGTGCTGACGACATTAAGATA 1408

Qy	175	GGGATTTATCTATTCAGAAAAGAGCTCGACATTTCTCTTAACCATGGATCTTAATCATC	1813
Db	1407	GGGATTTATCTATTCAGAAAAGAGCTCGAATATTTCTCTTAACCATGGATCTTAATCATC	1348
Qy	1814	TAAAGTTTCTCTTCAGAGATGCTCTGTGTGGGAGAGAGAAAGCCCTCAGCAAAAGTATG	1873
Db	1347	TAAAGTTTCTCTTCAGAGATGCTCTGTGTGGGAGAGAGAAAGCCCTCAGCAAAAGTATG	1288
Qy	1874	TACAAGACAAATCATCAGCTTCATGAGCCGACGAGTGGCGAAGAAATCTCTCCAGAGAAAGC	1933
Db	1287	TACAAGACAAATCATCAGCTTCATGAGCCGACGAGTGGCGAAGAAATCTCTCCAGAGAAAGC	1228
Qy	1934	GCATATTATATGTGTGTGAGATGCAAGAAATATGCGCCAGAGATGATCATGATGCCCTTG	1993
Db	1227	GCATATTATATGTGTGTGAGATGCAAGAAATATGCGCCAGAGATGATCATGATGCCCTTG	1168
Qy	1994	TGCAAAATATATAGCAAAAGGTGTGAGTGTGAAAACATGAGACATGAAAAACCTTGCGCA	2053
Db	1167	TGCAAAATATATAGCAAAAGGTGTGAGTGTGAAAACATGAGACATGAAAAACCTTGCGCA	1108
Qy	2054	CTTTTAAAGAGAAAAAGCTACTCTTCAGATATTTGGTCAATA	2097
Db	1107	CTTTTAAAGAGAAAAAGCTACTCTTCAGATATTTGGTCAATA	1064

	RESULT	9				
Locus	BD077780		390 bp	DNA	linear	PAT 27-AUG-2002
DEFINITION	5'EST of secretory protein in brain.					
ACCESSION	BD077780					
VERSION	BD077780.1		GI:22623383			
KEYWORDS	JP 2001512015-A/65.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrates; Euteleostomi;						
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.						
REFERENCE	1 (bases 1 to 390)					
AUTHORS	Edwards,J.B.D.M., Duclert,A. and Lacroix,B.					
TITLE	5'EST of secretory protein in brain					
JOURNAL	Patent : JP 2001512015-A 65 21-AUG-2001; GENBANK					
COMMENT	 OS Homo sapiens (human) PN JP 2001512015-A/65 PD 21-AUG-2001 PF 31-JUL-1998 JP 2000505293 PR 01-AUG-1997 US 08/905823 PI JEAN BAPTISTE DUMAS MILNE EDWARDS, AYMERIC DUCLERT, BRUNO PI LAEROIX  PC C12N15/09,C07K41/47,C12M1/00,C12P21/02,C12Q1/68,C12N15/00 CC Von Heljne matrix CC score 6..9 CC seq SLSPASHSHSVSC/SN FH Key Location/Qualifiers FT sig_peptide 289..357. Location/Qualifiers 1..390 "/organism='Homo sapiens' '/mol_type='genomic DNA" /db_xref='taxon:9606"					
FEATURES	Source					
ORIGIN						
Query Match	18.4%	Score 386;	DB 6;	Length 390;		
Best Local Similarity	100.0%;	Pred. NO. 4.*e-199;				
Matches 386;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;		
CY	970	GTACAAGCCTACTCAAGA CTGCGATTGAAGTAAAGA GACATGGTCCTTTTG	10259			
Db	3	GTA CAAGACCTA CTC AA AGA CTGCAGCTT GAAG TAAA GAGG AC GTGG TC TT TG	62			
CY	1030	AAAAATAAG GCAGAC ACA AAA GAGAG ATCA CCTAAC CCCAGCATAT ACTGC GGGA	10899			
Db	63	AAATTAAG GAGAGACA CAAGAAG AAAGAG CT TA CTT AACCCCAG CATAT ACTGC GGGA	122			

Qy	1090	TGTTCTCTCCAGTTCATTTTTTACTCTGGTGCTTGAATCCGACCAATTCCATAAAAGSCA	1149
Db	123	TGTTCTCTCCAGTTCATTTTTTACTCTGGTGCTTGAATCCGACCAATTCCATAAAAGSCA	182
Qy	1150	TTTTTGCGAGCCCTTGTGACTATACCAAGTGCAGTCTGTAAAAAGCAGGCTTACAGAG	1209
Db	183	TTTTTGCGAGCCCTTGTGACTATACCAAGTGCAGTCTGTAAAAAGCAGGCTTACAGAG	242
Qy	1210	CTGTGCACTAACCAAGGGGAGCGCATTAAGCGCTTGTATAGAATGTGCTGGCTGC	1269
Db	243	CTGTGCACTAACCAAGGGGAGCGCATTAAGCGCTTGTATAGAATGTGCTGGCTGC	302
Qy	1270	TTGTGTGATCTCCTCTCGCTTTCCTTCTTGTGCAGGCAACACTCACTCTCTGTGTGAA	1329
Db	303	TTGTGTGATCTCCTCCTCGCTTTCCTTCTTGTGCAGGCAACACTCACTCTCTGTGTGAA	362
Qy	1330	CATCTTCTTAACCTTGAACCCAGACC	1355
Db	363	CATCTTCTTAACCTTGAACCCAGACC	388

LOCUS	LOCUS	1353 bp	DNA	linear	PRI 14-DEC-1999
FI121202S04	FI121202S04				
DEFINITION	Homo sapiens methionine synthase reductase (MTRR) gene, exon 5.				
ACCESSION	AF121205				
VERSION	AF121205.1	GI:6572530			
KEYWORDS					
SEGMENT	4 of 12				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 1353) Leclerc,D., Odleyev,M., Wu,Q., Wilson,A., Huizenga,J.J., Rozen,R., Scherer,S.W. and Gravel,R.A.				
TITLE	Molecular cloning, expression and physical mapping of the human methionine synthase reductase gene				
JOURNAL	Gene 240 (1), 75-88 (1999)				
MEDLINE	20033550				
PUBMED	10564814				
REFERENCE	2 (bases 1 to 1353)				
AUTHORS	Leclerc,D.				
TITLE	Direct Submission				
JOURNAL	Submitted (20-JAN-1999) Human Genetics, Montreal Children's Hospital, 4060 Ste-Catherine West, Montreal, Quebec H3Z 2Z3, Canada				
FEATURES	Location/Qualifiers				
Source	1..1353				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
	/chromosome="5"				
exon	/map="p15.2-p15.3; between W1755 and D5S1957"				
	358..736				
	/gene="MTRR"				
	/number=5				
ORIGIN					
Query Match	18.2%; Score 381; DB 9; Length 1353;				
Best Local Similarity	100.0%; Pred. No. 2,2e-196;				
Matches 381; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				
Qy	401 GTTGAAGACTTGTCGTGAGCCGCGGATTCGTGACCTCTGACCGACCTCAGAAAGCATT	460			
Db	357 GTTGAAGACTTGTCGTGAGCCGCGGATTCGTGACCTCTGACCGACCTCAGAAAGCATT	416			
Qy	461 TTAGGTCAAGCAGACGACCAAGAGAGATTAAGTGGCGCACTCCCGTGGCATCACTGCAT	520			
Db	417 TTAGGTCAAGCAGACGACCAAGAGAGATTAAGTGGCGCACTCCCGTGGCATCACTGCAT	476			
Qy	521 CTTGAGACAGACCTTTGTGAAGTCAGAGCTGCTACACATTGAATCTTCAAGTGAAGCTTC	580			

Db 477 CCTTGAGAGACAGACCTTGTGAAGTCAGAGCTGCTACATGTAATCTCAAGTCGAGCTTC 536  
|||  
Qy 581 TGAGATTGATGATTCAGAGAAAGAAAGATTCTGAGCTTTTGAAGCAAAATGCAAGTGAACA 640  
|||  
Db 537 TGAGATTGATGATTCAGAGAAAGAAAGATTCTGAGCTTTTGAAGCAAAATGCAAGTGAACA 596  
|||  
Qy 641 GCAACCAATCCAAATGTTGAATTTGAAGACTTTGAGTCTCACTTACCCGTCGGTACCCC 700  
|||  
Db 597 GCAACCAATCCAAATGTTGAATTTGAAGACTTTGAGTCTCACTTACCCGTCGGTACCCC 656  
|||  
Qy 701 CACTCTCAGAGCTCTCTGTAATTTCTGTTTACCCCAAGAAATTTTACAGATGATC 760  
|||  
Db 657 CACTCTCAGAGCTCTCTGTAATTTCTGTTTACCCCAAGAAATTTTACAGATGATC 716  
|||  
Qy 761 TGCAGAGATCTCTGGCCAGG 781  
|||  
Db 717 TGCAGAGATCTCTGGCCAGG 737  
|||

## RESULT 11

AC010346

LOCUS AC010346 109626 bp DNA linear PRI 10-NOV-2000  
DEFINITION Homo sapiens chromosome 5 clone CITB-H1\_2018B2, complete sequence.  
AC010346  
AC010346 GI:11136705  
VERSION HTG.  
KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 109626)  
DOI Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Unpublished  
2 (bases 1 to 109626)  
REFERENCE  
AUTHORS DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
3 (bases 1 to 109626)  
REFERENCE  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Submitted (10-NOV-2000) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
On Nov 10, 2000 this sequence version replaced gi:9256196.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www-shgc.stanford.edu  
STS Content:  
Quality: Phrap Quality >=40 99.9% of Sequence;  
WI-9255 G05749.

## FEATURES

source

1. .109626  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/clone="CITB-H1\_2018B2"

## ORIGIN

Query Match 15.7%; Score 330; DB 9; Length 109626;  
Best Local Similarity 99.7%; Pred. No. 1.3e-168;  
Matches 380; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 401 GTTTAGAACTTGTGTTGAGCCGTGATTTCTGACTTGGCCAGCCCTCAGAAACATT 460  
|||  
Db 88571 GTTTAGAACTTGTGTTGAGCCGTGATTTCTGACTTGGCCAGCCCTCAGAAACATT 88630  
|||  
Qy 461 TTAGTCAAGCAGAGACAAAGAGATTAAGTGGCGCACTCCCGTGGCATCACTTCGAT 520  
|||  
Db 88631 TTAGTCAAGCAGAGACAAAGAGATTAAGTGGCGCACTCCCGTGGCATCACTTCGAT 88690  
|||

Qy 521 CCTTGAGAGACAGACCTTGTGAAGTCAGAGCTGCTACATGTAATCTCAAGTCGAGCTTC 580  
|||  
Db 88691 CCTTGAGAGACAGACCTTGTGAAGTCAGAGCTGCTACATGTAATCTCAAGTCGAGCTTC 88750  
|||  
Qy 581 TGAGATTGATGATTCAGAGAAAGAAAGATTCTGAGCTTTTGAAGCAAAATGCAAGTGAACA 640  
|||  
Db 88751 TGAGATTGATGATTCAGAGAAAGAAAGATTCTGAGCTTTTGAAGCAAAATGCAAGTGAACA 88810  
|||  
Qy 641 GCAACCAATCCAAATGTTGAATTTGAAGACTTTGAGTCTCACTTACCCGTCGGTACCCC 700  
|||  
Db 88811 GCAACCAATCCAAATGTTGAATTTGAAGACTTTGAGTCTCACTTACCCGTCGGTACCCC 88870  
|||  
Qy 701 CACTCTCAGAGCTCTCTGTAATTTCTGTTTACCCCAAGAAATTTTACAGATGATC 760  
|||  
Db 88871 CACTCTCAGAGCTCTCTGTAATTTCTGTTTACCCCAAGAAATTTTACAGATGATC 88930  
|||  
Qy 761 TGCAGAGATCTCTGGCCAGG 781  
|||  
Db 88931 TGCAGAGATCTCTGGCCAGG 88951  
|||

## RESULT 12

AC025174

LOCUS AC025174 110756 bp DNA linear PRI 28-MAR-2002  
DEFINITION Homo sapiens chromosome 5 clone CTD-2072124, complete sequence.  
AC025174  
AC025174 GI:19774456  
VERSION HTG.  
KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 110756)  
DOI Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Unpublished  
2 (bases 1 to 110756)  
REFERENCE  
AUTHORS DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (07-MAR-2000) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
3 (bases 1 to 110756)  
REFERENCE  
AUTHORS DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (07-MAR-2002) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
4 (bases 1 to 110756)  
REFERENCE  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Submitted (28-MAR-2002) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
On Mar 28, 2002 this sequence version replaced gi:19224767.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www-shgc.stanford.edu  
Quality: Phrap Quality >=40 100% of Sequence;  
Estimated Total Number of Errors is 0.

## FEATURES

source

1. .110756  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/clone="CTD-2072124"

## ORIGIN

Query Match 15.7%; Score 330; DB 9; Length 110756;  
Best Local Similarity 99.7%; Pred. No. 1.3e-168;  
Matches 380; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 401 GTTTAGAACTTGTGTTGAGCCGTGATTTCTGACTTGGCCAGCCCTCAGAAACATT 460  
|||

Db 20040 GTTTAGAACTTGTGTTGACCGCTGACTTCTGGAAGCTTGGCCAGCCCTCAGAAAGCATT 20099  
Qy 461 TTAGTCAACAGAGAGCAAGAGAGATTAAGTGGCGCACTCCGATGGCATCCTGCAT 520  
Db 20100 TTAGTCAACAGAGAGCAAGAGAGATTAAGTGGCGCACTCCGATGGCATCCTGCAT 20159  
Qy 521 CCTTGAAGAGACAGCTTGTGAAGTCAAGCTGCTACACATTGAATCTCAAGTGAAGCTTC 580  
Db 20160 CCTTGAAGAGACAGCTTGTGAAGTCAAGCTGCTACACATTGAATCTCAAGTGAAGCTTC 20219  
Qy 581 TGAGATTCAGATGATTGAGAAAGAGATTCGAGGTTTGAAGCAAAATGAGTGAACA 640  
Db 20220 TGAGATTCAGATGATTGAGAAAGAGATTCGAGGTTTGAAGCAAAATGAGTGAACA 20279  
Qy 641 GCAACCAATCAATGTTTAATTGAAGACATTTGAGTCCATTCACCTTCCGATACCC 700  
Db 20280 GCAACCAATCAATGTTTAATTGAAGACATTTGAGTCCATTCACCTTCCGATACCC 20339  
Qy 701 CACTCTCAACAGCTCTCTGAATTTCTGTGTTTACCCCAAAATTTTACAGTATC 760  
Db 20340 CACTCTCAACAGCTCTCTGAATTTCTGTGTTTACCCCAAAATTTTACAGTATC 20399  
Qy 761 TGCAGAGTCTCTTGGCCAG 781  
Db 20400 TGCAGAGTCTCTTGGCCAG 20420

RESULT 13  
AC022921 158199 bp DNA linear HTG 12-MAR-2000  
LOCUS Homo sapiens clone RP11-138P20, WORKING DRAFT SEQUENCE, 12  
DEFINITION Unordered pieces.  
AC022921  
AC022921.2 GI:7229868  
HTG; HTG PHASE1; HTGS\_DRAFT.  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 158199)  
Britten, B., Linton, L., Nussbaum, C. and Lander, E.  
Homo sapiens, clone RP11-138P20  
2 (bases 1 to 158199)  
Britten, B., Linton, L., Nussbaum, C., Lander, E., Abraham, H., Allen, N.,  
Anderson, S., Baldwin, J., Barna, N., Beckert, R., Bede, F.,  
Boguslavsky, L., Bouhgalter, B., Brown, A., Burkert, G., Casle, A.,  
Choedel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,  
Darellano, K., Dewar, K., Domino, M., Doyle, M., Fenebor, J.,  
Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J.,  
Gardyna, S., Grant, G., Hages, B., Heath, A., Horton, L.,  
Howard, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,  
Lander, T., Lebecky, J., Levine, R., Liu, C., Liu, G., Locke, K.,  
Macdonald, P., Margulis, N., McEwan, P., McGuck, A., McKernan, K.,  
McPherson, R., Meldrum, J., Meneses, L., Morrow, J., Naylor, J.,  
Norman, C. H., O'Connor, T., O'Donnell, P., Oliver, I. M., Peterson, K.,  
Pierre, N., Pisan, C., Pollard, V., Raymond, C., Riley, R., Rothman, D.,  
Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,  
Stojanovic, N., Subramanian, A., Talamas, J., Teefaye, S., Theodore, J.,  
Tirrell, A., Vasilev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,  
Zimmer, A. and Zody, M.

TITLE Direct Submission  
JOURNAL Submitted (07-FEB-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
COMMENT On Mar 12, 2000 this sequence version replaced gi:6921909.  
All repeats were identified using RepeatMasker:  
Smit, A.P.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information  
Center project name: L6314  
Center clone name: 138\_P20  
----- Summary Statistics  
Sequencing vector: M13; M77815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 152656 bases at least Q40  
Consensus quality: 155474 bases at least Q30  
Consensus quality: 156388 bases at least Q20  
Insert size: 178000; agarose-fp  
Insert size: 157099; sum-of-ctnigs  
Quality coverage: 4.4 in Q20 bases; agarose-fp  
Quality coverage: 5.0 in Q20 bases; sum-of-ctnigs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 12 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1 1283: contig of 1283 bp in length  
\* 1284 1383: gap of 100 bp  
\* 1384 4203: contig of 2820 bp in length  
\* 4204 4303: gap of 100 bp  
\* 4304 6786: contig of 2483 bp in length  
\* 6787 6886: gap of 100 bp  
\* 6887 9683: contig of 2797 bp in length  
\* 9684 9783: gap of 100 bp  
\* 9784 12902: contig of 3119 bp in length  
\* 12903 13002: gap of 100 bp  
\* 13003 16429: contig of 3427 bp in length  
\* 16430 16529: gap of 100 bp  
\* 16530 25201: contig of 8672 bp in length  
\* 25202 25301: gap of 100 bp  
\* 25302 36759: contig of 11458 bp in length  
\* 36760 36859: gap of 100 bp  
\* 36860 53921: contig of 17062 bp in length  
\* 53922 54021: gap of 100 bp  
\* 54022 72054: contig of 18033 bp in length  
\* 72055 72154: gap of 100 bp  
\* 72155 102527: contig of 30373 bp in length  
\* 102528 102627: gap of 100 bp  
\* 102628 158199: contig of 55572 bp in length.  
Location/Qualifiers  
1..158199  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="RP11-138P20"  
/clone\_lib="RP11-138P20"  
1..1283  
/note="assembly\_fragment"  
1384..4203  
/note="assembly\_fragment"  
4304..6786  
/note="assembly\_fragment"  
6887..9683  
/note="assembly\_fragment"  
clone\_end:SP6  
vector\_side:left  
/note="assembly\_fragment"  
9784..12902  
/note="assembly\_fragment"  
13003..16429  
/note="assembly\_fragment"  
16530..25201  
/note="assembly\_fragment"  
clone\_end:T7  
vector\_side:right  
25302..36759  
/note="assembly\_fragment"



```

misc_feature      36860..53921
                  /note="assembly_fragment"
misc_feature      54022..72054
                  /note="assembly_fragment"
misc_feature      72155..102527
                  /note="assembly_fragment"
misc_feature      102628..158199
                  /note="assembly_fragment"

```

## ORIGIN

Query Match 13.3%; Score 279; DB 2; Length 158199;

Best Local Similarity 99.5%; Pred. No. 1.1e-140; Matches 379; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 401 GTTTGAATCTGGTGGTGGCCGTGATCTGCACTCGGCGAGCCCTCAGAAAGCAT 460
    |||
DB 3446 GTTTGAATCTGGTGGTGGCCGTGATCTGCACTCGGCGAGCCCTCAGAAAGCAT 3505
    |||
QY 461 TTAGGTCAAGCAGAGCAGAGAGATAGTGGCGACCTCCGGTGCGATCCTGCAT 520
    |||
DB 3506 TTAGGTCAAGCAGAGCAGAGAGATAGTGGCGACCTCCGGTGCGATCCTGCAT 3565
    |||
QY 521 CCTTGAAGCAGACCTTGTGAAGTCAAGCTGCTACATTAATCTCAAGTCACTTC 580
    |||
DB 3566 CCTTGAAGCAGACCTTGTGAAGTCAAGCTGCTACATTAATCTCAAGTCACTTC 3625
    |||
QY 581 TGAAGTCAATGATTTAGAGAAAGATTTCTGAGTTTGAAGCAAAATGCAATGAACA 640
    |||
DB 3626 TGAAGTCAATGATTTAGAGAAAGATTTCTGAGTTTGAAGCAAAATGCAATGAACA 3685
    |||
QY 641 GCACCAATCCATGTTGTAATGAAGCTTGAAGCTTCACTTACCCGCTGGTACCCC 700
    |||
DB 3686 GCACCAATCCATGTTGTAATGAAGCTTGAAGCTTCACTTACCCGCTGGTACCCC 3745
    |||
QY 701 CACTCTCACAAGCCTCTCTGAATATTCCTGTTTACCCCGAATATTTACAGATCATC 760
    |||
DB 3746 CACTCTCACAAGCCTCTCTGAATATTCCTGTTTACCCCGAATATTTACAGATCATC 3805
    |||
QY 761 TGCAGAGTCTCTTGGCCAGG 781
    |||
DB 3806 TGCAGAGTCTCTTGGCCAGG 3826
    |||

```

RESULT 14  
AC021609  
LOCUS  
DEFINITION Homo sapiens clone RP11-259D10, WORKING DRAFT SEQUENCE, 6 unordered  
pieces.  
AC021609  
AC021609.3 GI:7230210  
VERSION  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE  
ORGANISM Homo sapiens

Bukaryovca; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
1 (bases 1 to 167237)  
Birren, B., Linton, L., Nuebaum, C., Lander, E., Abraham, H., Allen, N.,  
Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Bede, P.,  
Bogunlavsky, L., Boukhalter, B., Brown, A., Burkett, G., Castle, A.,  
Chapel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,  
DeArillano, K., Dewar, K., Domini, M., Doyle, M., Feneator, J.,  
Ferrera, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J.,  
Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
Howland, J. C., Johnson, R., Jones, C., Kann, L., Karataca, A., Klein, J.,  
Lander, T., Lehoczeky, J., Levine, R., Lieu, C., Liu, G., Locke, K.,  
Macdonald, P., Marquis, N., McMan, P., McGurk, A., McKernan, K.,  
McPheeters, R., Meldrum, J., Meneus, L., Morrow, J., Naylor, J.,  
Norman, C. H., O'Connor, T., O'Donnell, P., Oliver, T. M., Peterson, K.,  
Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,

ROY, A., Santos, R., Severy, P., Spencer, B., Strange-Thomann, N.,  
Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,  
Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,  
Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Mar 12, 2000 this sequence version replaced gi:6899697.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

## COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submission@genome.wi.mit.edu  
----- Project Information -----  
Center project name: L5818  
Center clone name: 259\_D\_10  
----- Summary Statistics -----  
Sequencing vector: M13; M77815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 161183 bases at least Q40  
Consensus quality: 164380 bases at least Q30  
Consensus quality: 165590 bases at least Q20  
Insert size: 164000; agarose-fp  
Insert size: 166737; sum-of-contigs  
Quality coverage: 5.1 in Q20 bases; agarose-fp  
Quality coverage: 5.0 in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 6 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1  
3657 3656: contig of 3656 bp in length  
3757 3756: gap of 100 bp  
9437 9436: contig of 5680 bp in length  
9537 9536: gap of 100 bp  
27769 27768: contig of 1832 bp in length  
27869 27868: gap of 100 bp  
52058 52058: contig of 24190 bp in length  
52059 52158: gap of 100 bp  
52159 80100: contig of 27942 bp in length  
80101 80200: gap of 100 bp  
80201 167237: contig of 87037 bp in length.

## FEATURES

source  
1..167237  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="RP11-259D10"  
/clone\_lib="RP11-259D10 Human Male BAC"  
1..3656  
/note="assembly\_fragment  
clone end:17  
vector side:left"  
misc\_feature  
3757..9436  
/note="assembly\_fragment"  
misc\_feature  
9537..27768  
/note="assembly\_fragment  
clone end:806  
vector side:right"  
misc\_feature  
27869..52058  
/note="assembly\_fragment"  
misc\_feature  
52159..80100  
/note="assembly\_fragment"  
misc\_feature  
80201..167237  
/note="assembly\_fragment"



## ORIGIN

Query Match 13.3%; Score 279; DB 2; Length 167237;  
Best Local Similarity 99.5%; Pred. No. 1.1e-140;  
Matches 379; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 401 GTTTAGAACTTGTGGTGGACCGCTGGATTGCTGACCTTGGCCAGCCCTCAGAAAGCATT 460  
DB 62083 GTTTAGAACTTGTGGTGGACCGCTGGATTGCTGACCTTGGCCAGCCCTCAGAAAGCATT 62142  
QY 461 TTAGAGTCAGACAGACAGACAGAGATAGTGGCGCACTCCGGTGGATCACTGACAT 520  
DB 62143 TTAGAGTCAGACAGACAGACAGAGATAGTGGCGCACTCCGGTGGATCACTGACAT 62202  
QY 521 CCTGAGACAGACAGCTTGTGAGTCAGAGCTGCTACACATTGATCTCAAGTCAGAGCTTC 580  
DB 62203 CCTGAGACAGACAGCTTGTGAGTCAGAGCTGCTACACATTGATCTCAAGTCAGAGCTTC 62262  
QY 581 TGAGATTCGATGATTCAGGAAAGAGATTCCTGAGCTTGGAGCAAAATGCAGTGAACA 640  
DB 62263 TGAGATTCGATGATTCAGGAAAGAGATTCCTGAGCTTGGAGCAAAATGCAGTGAACA 62322  
QY 641 GCAACCAATCAATGTTGTAATTAAGAGATCTTGAAGTCCGCTTACCCGTTCCGTTACCCC 700  
DB 62323 GCAACCAATCAATGTTGTAATTAAGAGATCTTGAAGTCCGCTTACCCGTTCCGTTACCCC 62382  
QY 701 CACTTCACAAGCCTCTCTGAATATTCCTGCTTACCCTCCAGATATTTACAGGTATC 760  
DB 62383 CACTTCACAAGCCTCTCTGAATATTCCTGCTTACCCTCCAGATATTTACAGGTATC 62442  
QY 761 TGCAGAGTCTCTTGGCCAG 781  
DB 62443 TGCAGAGTCTCTTGGCCAG 62463

## RESULT 15

AC091945 177596 bp DNA linear HTG 09-JUN-2001  
LOCUS Homo sapiens chromosome 5 clone RP11-35616, WORKING DRAFT SEQUENCE,  
DEFINITION 27 unordered pieces.

AC091945 AC091945.1 GI:14333881  
VERSION HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 177596)

REFERENCE DOB Joint Genome Institute.  
AUTHORS Unpublished  
TITLE DOB Joint Genome Institute.  
JOURNAL 2 (bases 1 to 177596)

REFERENCE DOB Joint Genome Institute.  
AUTHORS Direct Submission.  
TITLE Submitted (09-JUN-2001) Production Sequencing Facility, DOE Joint  
JOURNAL Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

COMMENT -----Genome Center  
Center: Joint Genome Institute  
Center Code: JGI  
Web site: http://www.jgi.doe.gov  
-----

Project Information  
Center Project Name: 543267  
Center clone name: RPCT-11\_35616  
-----

Summary Statistics  
Consensus quality: 152700 bases at least Q40  
Consensus quality: 168451 bases at least Q30  
Consensus quality: 170419 bases at least Q20  
Estimated insert size: 226080; agarose-fp estimation  
Estimated insert size: 174996; sum-of-contigs estimation  
Quality coverage: 7.75 in Q20 bases; agarose-fp estimation  
Quality coverage: 10.01 in Q20 bases; sum-of-contigs estimation.

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 27 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1207: contig of 1207 bp in length  
1208 1307: gap of unknown length  
1308 3385: contig of 2078 bp in length  
3386 3386: gap of unknown length  
3486 5553: contig of 2068 bp in length  
5554 5653: gap of unknown length  
5654 7030: contig of 1377 bp in length  
7031 7130: gap of unknown length  
7131 9069: contig of 1939 bp in length  
9070 9169: gap of unknown length  
9170 11452: contig of 2283 bp in length  
11453 11552: gap of unknown length  
11553 16002: contig of 4450 bp in length  
16003 16102: gap of unknown length  
16103 21083: contig of 4981 bp in length  
21084 21883: gap of unknown length  
21884 26893: contig of 5710 bp in length  
26894 26993: gap of unknown length  
26993 29987: contig of 2994 bp in length  
29988 32949: gap of unknown length  
32949 33049: gap of unknown length  
33050 38757: contig of 5708 bp in length  
38758 38857: gap of unknown length  
38857 45202: contig of 6345 bp in length  
45203 45302: gap of unknown length  
45303 50911: contig of 5609 bp in length  
50912 51011: gap of unknown length  
51012 55880: contig of 4869 bp in length  
55881 55980: gap of unknown length  
55981 63189: contig of 7209 bp in length  
63190 63289: gap of unknown length  
63290 66632: contig of 6343 bp in length  
66633 69732: gap of unknown length  
69733 76377: contig of 6645 bp in length  
76378 76477: gap of unknown length  
76477 83051: contig of 6574 bp in length  
83052 83151: gap of unknown length  
83152 94929: contig of 11778 bp in length  
94930 95029: gap of unknown length  
95030 104346: contig of 9317 bp in length  
104347 104446: gap of unknown length  
104447 115275: contig of 10829 bp in length  
115276 115375: gap of unknown length  
115376 123488: contig of 8113 bp in length  
123489 123588: gap of unknown length  
123589 133172: contig of 9584 bp in length  
133173 133272: gap of unknown length  
133273 145495: contig of 12223 bp in length  
145496 145595: gap of unknown length  
145596 160270: contig of 14675 bp in length  
160271 160370: gap of unknown length  
160371 177596: contig of 17226 bp in length.

## FEATURES

## source

1. 177596  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/clone="RP11-35616"  
/clone\_11b="RPCT human BAC library 11"

## ORIGIN

Query Match 13.3%; Score 279; DB 2; Length 177596;  
Best Local Similarity 99.5%; Pred. No. 1.1e-140;

Matches 379; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
OY 401 GTTGAACCTTGTGGTGGCCGTGGATCTGGCCAGCCCTCAGAAAGCAT 460
    |||||
Db 83524 GTTGAACCTTGTGGTGGCCGTGGATCTGGCCAGCCCTCAGAAAGCAT 83583
    |||||
OY 461 TTAGGTCAAGCAGAGACAAAGAGAGATAGTGGCGCACTCCCGGTGGCATCCTGCAT 520
    |||||
Db 83584 TTAGGTCAAGCAGAGACAAAGAGAGATAGTGGCGCACTCCCGGTGGCATCCTGCAT 83643
    |||||
OY 521 CCTGAGAGACAGACCTTGTGAAGTCAGAGCTGCTACACATTGAATCTCAAGTCAGCTTC 580
    |||||
Db 83644 CCTGAGAGACAGACCTTGTGAAGTCAGAGCTGCTACACATTGAATCTCAAGTCAGCTTC 83703
    |||||
OY 581 TGAGATTGATGATTCAGAGAAAGAAAGATTCTGAGGTTTGAAGCAAAATGCAAGTGAACA 640
    |||||
Db 83704 TGAGATTGATGATTCAGAGAAAGAAAGATTCTGAGGTTTGAAGCAAAATGCAAGTGAACA 83763
    |||||
OY 641 GCAACCAATCCATGTTGTAATTGAAGACTTTGAGTCTCACTTACCCGTTGCTACCCC 700
    |||||
Db 83764 GCAACCAATCCATGTTGTAATTGAAGACTTTGAGTCTCACTTACCCGTTGCTACCCC 83823
    |||||
OY 701 CACTCTCACAGCCTCTCTGAATTAATCTGTGTTTACCCCCAGAAATTTTACAGGTACATC 760
    |||||
Db 83824 CACTCTCACAGCCTCTCTGAATTAATCTGTGTTTACCCCCAGAAATTTTACAGGTACATC 83883
    |||||
OY 761 TGCAGAGTCTCTTGGCCAGG 781
    |||||
Db 83884 TGCAGAGTCTCTTGGCCAGG 83904
    |||||
```

Search completed: August 27, 2005, 09:38:56  
Job time : 6010.21 secs

**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 13:32:20 ; Search time 733.49 Seconds  
(without alignments)  
16924.161 Million cell updates/sec

Title: US-09-371-347A-43

Perfect score: 2097  
Sequence: 1 atgagagaggttcctgact.....ttcagatatttgcataaa 2097

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*  
1: geneseqn19808:\*  
2: geneseqn19908:\*  
3: geneseqn20008:\*  
4: geneseqn20018:\*  
5: geneseqn20028:\*  
6: geneseqn20038:\*  
7: geneseqn20048:\*  
8: geneseqn20058:\*  
9: geneseqn20068:\*  
10: geneseqn20078:\*  
11: geneseqn20088:\*  
12: geneseqn20098:\*  
13: geneseqn20108:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2094	99.9	2094	11	ADM43212	Adm43212 Human met
2	2046	97.6	3259	5	AA65070	AA65070 DNA encod
3	2046	97.6	3259	5	AA691226	AA691226 Human ech
4	2046	97.6	3259	11	ADM43206	Adm43206 Human full
5	2043	97.4	2094	11	ADM43208	Adm43208 Human met
6	1992	95.0	2094	11	ADM43209	Adm43209 Human met
7	1944	92.7	3259	3	AA58935	AA58935 DNA encod
8	1893	90.3	3270	13	ADQ87538	Adq87538 Human tum
9	1800	85.8	2091	11	ADM43216	Adm43216 Human met
10	1800	85.8	2091	11	ADM43214	Adm43214 Human met
11	1701	81.1	3256	3	AA58977	AA58977 A human m
12	1640	78.2	3255	3	AA58976	AA58976 A human m
13	1640	78.2	3256	3	AAQ39029	Adq39029 Human SNP
14	1018	48.5	3274	13	ADQ39030	Adq39030 Human SNP
15	905	43.2	3189	13	ACH42470	Ach42470 Human dia
16	805	38.4	1986	4	AA541064	AA541064 cDNA enco
17	427	20.4	1663	4	AA541602	AA541602 cDNA enco
18	386	18.4	390	2	AA551820	AA551820 Human sec
19	330	15.7	591	2	ACH73174	ACH73174 Human gen
20	328	15.6	379	12	ACH86905	ACH86905 Human gen

21	279	13.3	591	12	ACH86540	ACH86540 Human gen
22	277	13.2	379	12	ACH82240	ACH82240 Human gen
23	225	10.7	503	5	Aa655069	Aa655069 DNA encod
24	188	9.0	525	12	ACH67438	ACH67438 Human gen
25	175	8.3	175	12	ACH81143	ACH81143 Human gen
26	158	7.5	2475	6	AD32365	AD32365 Human lun
27	158	7.5	2475	13	AD161720	AD161720 Human CDN
28	137	6.5	525	12	ACH73117	ACH73117 Human gen
29	124	5.9	175	12	ACH86848	ACH86848 Human gen
30	69	3.3	244	3	AA242736	AA242736 Human 5'
31	60	2.9	60	6	ABN36264	ABN36264 Human sp1
32	51	2.4	51	4	AA178548	AA178548 Human s11
33	30	1.4	1681	11	AD131127	AD131127 Human CDN
34	27	1.3	1835	5	Aa655071	Aa655071 DNA encod
35	26	1.2	26	3	Aa58955	Aa58955 PCR prime
36	26	1.2	26	3	Aa58939	Aa58939 PCR prime
37	26	1.2	26	3	ABX09549	ABX09549 Arteriosc
38	26	1.2	26	6	AA143713	AA143713 Pregestat
39	26	1.2	26	11	ADm43205	ADm43205 Human met
40	26	1.2	26	11	ADm43189	ADm43189 Human met
41	25	1.2	25	3	Aa58952	Aa58952 PCR prime
42	25	1.2	25	3	Aa58937	Aa58937 PCR prime
43	25	1.2	25	3	Aa58947	Aa58947 PCR prime
44	25	1.2	25	11	ADm43187	ADm43187 Human met
45	25	1.2	25	11	ADm43202	ADm43202 Human met

## ALIGNMENTS

RESULT 1	ADm43212	standard; cDNA; 2094 BP.
ID	ADm43212	standard; cDNA; 2094 BP.
XX	ADm43212;	
AC	ADm43212;	
XX	03-JUN-2004	(first entry)
DT	03-JUN-2004	(first entry)
XX	Human methionine synthase reductase CDS GI104 variant.	
DE	Human; ss; Methionine synthase reductase polypeptide; HsMTRR; cancer;	
XX	cardiovascular disease; neural tube defect; hyperhomocysteinemia;	
KW	chromosome 5p15.2-p15.3; SNP; single nucleotide polymorphism.	
KW	chromosome 5p15.2-p15.3; SNP; single nucleotide polymorphism.	
XX	Homo sapiens.	
OS	Homo sapiens.	
XX	Key	Location/Qualifiers
FT	CDS	1..2094
FT		/*tag= a
FT		/product= "HsMTRR"
FT		/partial
FT		/note= "No stop codon shown"
FT	variation	replace(66, A)
FT		/*tag= b
FT		/standard_name= "Single_nucleotide_polymorphism"
FT	variation	replace(110, G)
FT		/*tag= c
FT		/standard_name= "Single_nucleotide_polymorphism"
XX	US2003082676-A1.	
PN	US2003082676-A1.	
XX	01-MAY-2003.	
PD	01-MAY-2003.	
XX	10-AUG-1999;	99US-00371347.
XX	10-AUG-1999;	99US-00371347.
PR	16-JAN-1998;	98US-0071622P.
PR	16-JAN-1998;	98US-0071622P.
XX	15-JAN-1999;	99US-00232028.
XX	15-JAN-1999;	99US-00232028.
XX	(GRAY/) GRAVEL R. A.	
PA	(ROZE/) ROZEN R.	
PA	(LECL/) LECLERC D.	
PA	(WILS/) WILSON A.	
PA	(ROSE/) ROSENBLATT D.	



DB 1561 TCATCTCTCTGGAACAACAAATCTTTCACCTACAGATGACCCCTCAATCCCCATC 1620  
QY 1621 ATTAATGTTGGTCCAGAAACCGGCATAGCCCTTTATTTGGTTCCTACACATTAAGAG 1680  
DB 1621 ATTAATGTTGGTCCAGAAACCGGCATAGCCCTTTATTTGGTTCCTACACATTAAGAG 1680  
QY 1681 AAATCTCAAGAACACACACAGATGGAATTTTGGAGCATGTTGTTTGGCTGC 1740  
DB 1681 AAATCTCAAGAACACACACAGATGGAATTTTGGAGCATGTTGTTTGGCTGC 1740  
QY 1741 AGGCATTAAGATTAAGGATTAATCTATTCAAGAAAGCTCAGACATTTCTTAAGCATGAG 1800  
DB 1741 AGGCATTAAGATTAAGGATTAATCTATTCAAGAAAGCTCAGACATTTCTTAAGCATGAG 1800  
QY 1801 ATCTTAACATCATTAAGGTTTCTCTCAAGAGATGCTCTTGGGAGAGAGAGCC 1860  
DB 1801 ATCTTAACATCATTAAGGTTTCTCTCAAGAGATGCTCTTGGGAGAGAGAGCC 1860  
QY 1861 CCAAGCAAGATTAAGTACAGACACATCCAGCTTCATGGCAGAGGTTGGCAGATCTTC 1920  
DB 1861 CCAAGCAAGATTAAGTACAGACACATCCAGCTTCATGGCAGAGGTTGGCAGATCTTC 1920  
QY 1921 CTCCAGAGAACCGCCATATTTATGTGTGTGAGATGCAAGATATGGCCAGATGTA 1980  
DB 1921 CTCCAGAGAACCGCCATATTTATGTGTGTGAGATGCAAGATATGGCCAGATGTA 1980  
QY 1981 CATGATGCTCTTGTGCAATATTAAGCAAGAGCTTGAGTTGAAAACTAAGCATAG 2040  
DB 1981 CATGATGCTCTTGTGCAATATTAAGCAAGAGCTTGAGTTGAAAACTAAGCATAG 2040  
QY 2041 AAAACCTGGCCACTTAAAGAAAGAAACGCTACCTTCAGGATTTTGGTCA 2094  
DB 2041 AAAACCTGGCCACTTAAAGAAAGAAACGCTACCTTCAGGATTTTGGTCA 2094

RESULT 2  
AAS65070  
ID AAS65070 standard; cDNA; 3259 BP.  
XX AAS65070;  
AC AAS65070;  
XX 13-FEB-2002 (first entry)  
DT 13-FEB-2002 (first entry)  
XX  
DE DNA encoding novel human diagnostic protein #874.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001MO-US008631.  
XX  
PR 31-MAR-2000; 2000US-00540217.  
PR 23-AUG-2000; 2000US-00649167.  
XX  
PA (HYSB-) HYSB INC.  
XX  
PI Drmanac RT, Liu C, Tang YT;  
XX  
DR MPI; 2001-639362/73.  
DR P-PSDB; ABG00883.  
XX  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
PS Claim 1; SEQ ID NO 874; 103bp; English.

XX  
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probe, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (II) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX  
SQ Sequence 3259 BP; 944 A; 695 C; 674 G; 946 T; 0 U; 0 Other;

Query Match 97.6%; Score 2046; DB 5; Length 3259;  
Best Local Similarity 100.0%; Pred No. 0;

Matches 2086; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGAGAGTTCTGTATATATGCTTACACAGCAGAGGACAGGCAAGCCATGAGAA 60  
DB 80 ATGAGAGAGTTCTGTATATATGCTTACACAGCAGAGGACAGGCAAGCCATGAGAA 139  
QY 61 GAAATGTGTGAGCAAGCTGTGTACATGATTTTCTGCAGATTTTCACTATATAGTGA 120  
DB 140 GAAATGTGTGAGCAAGCTGTGTACATGATTTTCTGCAGATTTTCACTATATAGTGA 199  
QY 121 TCCGATTAAGTATGACCTTAAACCCGAAACAGCTCTCTGTGTGTGTTTCTACACG 180  
DB 200 TCCGATTAAGTATGACCTTAAACCCGAAACAGCTCTCTGTGTGTGTTTCTACACG 259  
QY 181 GGCACCGAGAGCCACCCGACACAGCCCGCAAGTTTGTAAAGAAATACAAACA 240  
DB 260 GGCACCGAGAGCCACCCGACACAGCCCGCAAGTTTGTAAAGAAATACAAACA 319  
QY 241 CTGCGGTTGATTTCTTGTCTACCTGCGGTATGAGTTTCTGCGTTCTGCGTTTCA 300  
DB 320 CTGCGGTTGATTTCTTGTCTACCTGCGGTATGAGTTTCTGCGTTCTGCGTTTCA 379  
QY 301 TACACCTACTTTTGGCAATGGGGGGAAGATTAATGAATTAACGACTTCAAGGCTTGA 360  
DB 380 TACACCTACTTTTGGCAATGGGGGGAAGATTAATGAATTAACGACTTCAAGGCTTGA 439  
QY 361 CGGCATTTCTATGACACTGACATGATGATCTGTAGTTTGAACCTTGTGTGAG 420  
DB 440 CGGCATTTCTATGACACTGACATGATGATCTGTAGTTTGAACCTTGTGTGAG 499  
QY 421 CCGTGAATGCTGGAATCTGCGCCAGCTTCAAGAACATTTTATGTTCAAGAGACAA 480  
DB 500 CCGTGAATGCTGGAATCTGCGCCAGCTTCAAGAACATTTTATGTTCAAGAGACAA 559  
QY 481 GAGGATTAAGTGGCGCACTCCCGTGGGATACCTGCAATCTTGAAGACAGACCTT 540  
DB 560 GAGGATTAAGTGGCGCACTCCCGTGGGATACCTGCAATCTTGAAGACAGACCTT 619  
QY 541 AAGTCAGAGCTGCTACATTAATCTCAAGTCAAGCTTCTGAGATTCTGATTCAGGA 600  
DB 620 AAGTCAGAGCTGCTACATTAATCTCAAGTCAAGCTTCTGAGATTCTGATTCAGGA 679  
QY 601 AGAAGGATTTGAGGTTTGAAGCAAAATGACATGACAGCAACCAATCCATGTTTGA 660  
DB 680 AGAAGGATTTGAGGTTTGAAGCAAAATGACATGACAGCAACCAATCCATGTTTGA 739

```

QY 661 ATTGAAGACTTTGAGTCTCACTTACCCGTTGGTACCCCACTCTCAAGCCTCTCTG 720
Db 740 ATTGAAGACTTTGAGTCTCACTTACCCGTTGGTACCCCACTCTCAAGCCTCTCTG 799
QY 721 AATATTCCTGTTTACCCCGAGATATTTACAGTACATCTGAGGAGTCTCTTGGCAG 780
Db 800 AATATTCCTGTTTACCCCGAGATATTTACAGTACATCTGAGGAGTCTCTTGGCAG 859
QY 781 GAGAAAGCCAGATATCTGACTTCAGAGATCCAGTCTTTCAGAGTCCCAATTTCAAG 840
Db 860 GAGAAAGCCAGATATCTGACTTCAGAGATCCAGTCTTTCAGAGTCCCAATTTCAAG 919
QY 841 GCGATTCAATTAATCAAGATGATGCAATTAACCACTCTGCTGGTGAATTTGACATT 900
Db 920 GCGATTCAATTAATCAAGATGATGCAATTAACCACTCTGCTGGTGAATTTGACATT 979
QY 901 TCAAAATCAGACTTTCTCTACAGCCTGAGAGATGCTGAGCGTATCTGCTCAACAT 960
Db 980 TCAAAATCAGACTTTCTCTACAGCCTGAGAGATGCTGAGCGTATCTGCTCAACAT 1039
QY 961 GATTCTGAGGTACAAAGCCTTCTCAAGA CTGAGCTTGAAGATTAAGAGACACTGC 1020
Db 1040 GATTCTGAGGTACAAAGCCTTCTCAAGA CTGAGCTTGAAGATTAAGAGACACTGC 1099
QY 1021 GTCTTTGAAAATTAAGGAGACACAAAGAAAGAGACTTACCTTACCCAGCATATA 1080
Db 1100 GTCTTTGAAAATTAAGGAGACACAAAGAAAGAGACTTACCTTACCCAGCATATA 1159
QY 1081 CTGCGGGAGATGTTCTCTCCAGTTCATTTTACCGGTGTCTGAAATTCGAGCAATTCCT 1140
Db 1160 CTGCGGGAGATGTTCTCTCCAGTTCATTTTACCGGTGTCTGAAATTCGAGCAATTCCT 1219
QY 1141 AAAAGGCAATTTTTCGAGCCCTTGTGACTATACAGTACAGTGTCTGAAAAGCGCAGG 1200
Db 1220 AAAAGGCAATTTTTCGAGCCCTTGTGACTATACAGTACAGTGTCTGAAAAGCGCAGG 1279
QY 1201 CTACAGAGAGCTGTGCACTTAACAAAGGAGCGGATTAATAGCGCTTTTACAGATGCC 1260
Db 1280 CTACAGAGAGCTGTGCACTTAACAAAGGAGCGGATTAATAGCGCTTTTACAGATGCC 1339
QY 1261 TGTGCTGCTGTTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1320
Db 1340 TGTGCTGCTGTTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1399
QY 1321 CTGCTGCAACATCTTCTTAACTTCAACCCAGACATATTCGTGTGCAAGCTCAAGTTTA 1380
Db 1400 CTGCTGCAACATCTTCTTAACTTCAACCCAGACATATTCGTGTGCAAGCTCAAGTTTA 1459
QY 1381 TTTTCAACCAAGAAAGCTCCATTTTGTCTTCAACATTTGTGAAATTTCTGTCTACCTGCA 1440
Db 1460 TTTTCAACCAAGAAAGCTCCATTTTGTCTTCAACATTTGTGAAATTTCTGTCTACCTGCA 1519
QY 1441 ACAGAGGTTCTGCGGAGGAGTATGTACAGGCTGCGCTTGTGTTGGTTCAGTT 1500
Db 1520 ACAGAGGTTCTGCGGAGGAGTATGTACAGGCTGCGCTTGTGTTGGTTCAGTT 1579
QY 1501 CTTCAGCCAAACATACATGATCCCATGAAGACGCGGAAAGCCCTGCTCTTAAGATA 1560
Db 1580 CTTCAGCCAAACATACATGATCCCATGAAGACGCGGAAAGCCCTGCTCTTAAGATA 1639
QY 1561 TCCATCTCTCTCTGAAACAAATTTCTTTCATCTTACCAAGATGACCCCTCAATCCCATC 1620
Db 1640 TCCATCTCTCTCTGAAACAAATTTCTTTCATCTTACCAAGATGACCCCTCAATCCCATC 1699
QY 1621 ATATATGTGGGTTCAGAGACCGGATAGCCCGTTTATTTGGGTCTTACAAATAGAGAG 1680
Db 1700 ATATATGTGGGTTCAGAGACCGGATAGCCCGTTTATTTGGGTCTTACAAATAGAGAG 1759
QY 1681 AAATCCAAAGAACCAACCCAGATGAATTTTGAAGCAATGTGGTGTGTTTGGCTGC 1740
Db 1760 AAATCCAAAGAACCAACCCAGATGAATTTTGAAGCAATGTGGTGTGTTTGGCTGC 1819
QY 1741 AGGCAATAGGATGAGATTAATCTATTCAAGAAAGAGCTCAGACATTTCTTAAGCATGG 1800

```

```

Db 1820 AGGCAATAGGATGAGATTAATCTATTCAAGAAAGAGCTCAGACATTTCTTAAGCATGG 1879
QY 1801 ATCTTAATCATCTTAAAGGTTTCTTCTCAGAGATGCTCTGTTGGGAGAGGAAGCC 1860
Db 1880 ATCTTAATCATCTTAAAGGTTTCTTCTCAGAGATGCTCTGTTGGGAGAGGAAGCC 1939
QY 1861 CCAGCAAGATATGTACAGACAAATCCAGCTTCATGGCCAGAGAGTGGCAGAAATCTTC 1920
Db 1940 CCAGCAAGATATGTACAGACAAATCCAGCTTCATGGCCAGAGAGTGGCAGAAATCTTC 1999
QY 1921 CTTCAGAGAGAGGCGCATATTTATGTGTGTGAGATGCAAGAAATATGGCCAGAGATGTA 1980
Db 2000 CTTCAGAGAGAGGCGCATATTTATGTGTGTGAGATGCAAGAAATATGGCCAGAGATGTA 2059
QY 1981 CATGATGCCCTTGTGCAAAATATATACCAAGAGGTTGAGATTGAAAACTAGAAAGCAATG 2040
Db 2060 CATGATGCCCTTGTGCAAAATATATACCAAGAGGTTGAGATTGAAAACTAGAAAGCAATG 2119
QY 2041 AAAACCTTGCCACTTTAAAGAAAGAAAACGCTACCTTCAGATATTTGGTCAATA 2097
Db 2120 AAAACCTTGCCACTTTAAAGAAAGAAAACGCTACCTTCAGATATTTGGTCAATA 2176

```

```

RESULT 3
AAC91226
ID AAC91226 standard; DNA; 3259 BP.
XX
AC AAC91226;
XX
DT 20-MAR-2001 (first entry)
XX
XX Human schizophrenia related gene SEQ ID NO: 23.
DE
XX
XX Human: schizophrenia; developmental disorder; spina bifida cystica;
KM Tourette's syndrome; bipolar illness; autism; conduct disorder;
KM attention deficit disorder; obsessive compulsive disorder;
KM chronic multiple tic syndrome; learning disorder; polymorphism; ds.
XX
OS Homo sapiens.
XX
PN M0200071754-A1.
XX
PD 30-NOV-2000.
XX
PF 24-MAY-2000; 2000MO-US014354.
XX
PR 25-MAY-1999; 99US-00318448.
XX
PA (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.
XX
PI Johnson WG, Stenroos ES;
XX
DR WPI; 2001-025174/03.
XX
XX
PT Diagnosing a developmental disorder, e.g. schizophrenia, by forming
PT datasets (DS) of genetic (e.g. genotypes of folate metabolism alleles)
PT and environmental variables affecting an individual and then comparing
PT these DS with reference DS.
XX
PS Disclosure; Page 142-143; 156pp; English.
XX
XX The present invention provides a novel method of estimating the
XX susceptibility of an individual to a developmental disorder using genetic
XX and environmental variables. The method can be used in the diagnosis,
XX prevention and treatment of disorders such as schizophrenia, spina bifida
XX cystica, Tourette's syndrome, bipolar illness, autism, conduct disorders,
XX attention deficit disorder, obsessive compulsive disorder, chronic
XX multiple tic syndrome and learning disorders such as dyslexia
XX
SQ Sequence 3259 BP; 944 A; 695 C; 674 G; 946 T; 0 U; 0 Other;

```

Query Match 97.6%; Score 2046; DB 5; Length 3259;



Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2096; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGAGAGAGTTTCTGTACTATATATGCTACACAGCAGGACAGGCAAAAGCCATCCGAGAA 60
Db 80 ATGAGAGAGTTTCTGTACTATATATGCTACACAGCAGGACAGGCAAAAGCCATCCGAGAA 139
QY 61 GAAATGTGAGACAGCTGTGATCATGATTTCTGAGATCTTCACTATATTAGTAA 120
Db 140 GAAATGTGAGACAGCTGTGATCATGATTTCTGAGATCTTCACTATATTAGTAA 199
QY 121 TCCGATTAAGTATGACCTTAAACCGAAGACGCTCTTGTGTGTGTGTTCTACACG 180
Db 200 TCCGATTAAGTATGACCTTAAACCGAAGACGCTCTTGTGTGTGTGTTCTACACG 259
QY 181 GGCACCGGAGACCCACCCGACACAGCCGCAAGTTGTTAAAGAAATACAGAACCAAC 240
Db 260 GGCACCGGAGACCCACCCGACACAGCCGCAAGTTGTTAAAGAAATACAGAACCAAC 319
QY 241 CTGCCGGTGAATTTCTTGTCTACCTGCGGTATGGGTACTGGGCTCGGTGATTCAGAA 300
Db 320 CTGCCGGTGAATTTCTTGTCTACCTGCGGTATGGGTACTGGGCTCGGTGATTCAGAA 379
QY 301 TACACCTACTTTTGCATGTGGGGGAGATTAATTGATTAACGACTTCAGAGCTTGAGCC 360
Db 380 TACACCTACTTTTGCATGTGGGGGAGATTAATTGATTAACGACTTCAGAGCTTGAGCC 439
QY 361 CGGCAATTTCTATGACACTGACATGACATGATGATGATGATTTAAGATTGATGATG 420
Db 440 CGGCAATTTCTATGACACTGACATGACATGATGATGATGATTTAAGATTGATGATG 499
QY 421 CGGTGATTTGCTGGACTCTGGCCAGGCCCTCAGAAAGCATTTTATAGTCAAGAGAGCA 480
Db 500 CGGTGATTTGCTGGACTCTGGCCAGGCCCTCAGAAAGCATTTTATAGTCAAGAGAGCA 559
QY 481 GAGGAGATTAAGTGGCGCACTCCGGTGGATCACTGCACTCTTGAAGACAGACCTTGTG 540
Db 560 GAGGAGATTAAGTGGCGCACTCCGGTGGATCACTGCACTCTTGAAGACAGACCTTGTG 619
QY 541 AAGTCAGAGCTGTACACATTGAATTCATAGTCGAGCTTCTGAGATTCCATGATTCAGA 600
Db 620 AAGTCAGAGCTGTACACATTGAATTCATAGTCGAGCTTCTGAGATTCCATGATTCAGA 679
QY 601 AGAAGAGATTCTGAGGTTTGAAGCAAAATGACGTAACAGCAACCAATCCAAATGTTGA 660
Db 680 AGAAGAGATTCTGAGGTTTGAAGCAAAATGACGTAACAGCAACCAATCCAAATGTTGA 739
QY 661 ATTGAAGACTTGAAGTCTCACTTACCCGTTCCGTTACCCCACTCTCAGAGCCTCTG 720
Db 740 ATTGAAGACTTGAAGTCTCACTTACCCGTTCCGTTACCCCACTCTCAGAGCCTCTG 799
QY 721 AATATTCCTGTTTACCCCAAGATATTTACAGATACATCTGAGAGATCTTGGCCAG 780
Db 800 AATATTCCTGTTTACCCCAAGATATTTACAGATACATCTGAGAGATCTTGGCCAG 859
QY 781 GAGGAAGCCAGATCTGTGACTTGAAGATCCAGTTTTCATGTGCAATTTCAAG 840
Db 860 GAGGAAGCCAGATCTGTGACTTGAAGATCCAGTTTTCATGTGCAATTTCAAG 919
QY 841 GCAGTTCAACTTACTAGATGATGCAATTAACCACTCTGCTGATGAATTTGACATT 900
Db 920 GCAGTTCAACTTACTAGATGATGCAATTAACCACTCTGCTGATGAATTTGACATT 979
QY 901 TCATAATACAGACTTTTCTATACGCTGAGATGCTTCAAGCTGATCTGCCCTAACAGT 960
Db 980 TCATAATACAGACTTTTCTATACGCTGAGATGCTTCAAGCTGATCTGCCCTAACAGT 1039
QY 961 GATTCTGAGGTACAAAGCCTTCCAAAGCTGAGATTGAAGATTAAGAGAGAGCACTGC 1020
Db 1040 GATTCTGAGGTACAAAGCCTTCCAAAGCTGAGATTGAAGATTAAGAGAGAGCACTGC 1099
QY 1021 GTCCCTTTGAAATAAAGCAGACACAAGAAAGAAAGAGCTACTTACCCAGCATATA 1080
```

```
Db 1100 GTCCCTTTGAAATAAAGCAGACACAAGAAAGAAAGAGCTACTTACCCAGCATATA 1159
QY 1081 CCGGGGGAGTGTCTCTCAGATTCAATTTTACCTGATCTTGAATCCGAGCATTTCT 1140
Db 1160 CCGGGGGAGTGTCTCTCAGATTCAATTTTACCTGATCTTGAATCCGAGCATTTCT 1219
QY 1141 AAAAAGCATTTTGTGAGCCCTTGTGACATTAACAGTACAGTCTGAAAAAGCAG 1200
Db 1220 AAAAAGCATTTTGTGAGCCCTTGTGACATTAACAGTACAGTCTGAAAAAGCAG 1279
QY 1201 CTACAGAGCTGTGACATTAACAGAGGAGCCGATTAATAGCCCTTGTGACAGATGC 1260
Db 1280 CTACAGAGCTGTGACATTAACAGAGGAGCCGATTAATAGCCCTTGTGACAGATGC 1339
QY 1261 TGTGCTGTGTGTGATCTCTCCGCTTCCCTGTCAGGCAACCACTCAGTCTC 1320
Db 1340 TGTGCTGTGTGTGATCTCTCCGCTTCCCTGTCAGGCAACCACTCAGTCTC 1399
QY 1321 CTGCTGAAATCTTCTTAACTTCAACCCAGACATATTCGTGTCAGTCAAGTTTA 1380
Db 1400 CTGCTGAAATCTTCTTAACTTCAACCCAGACATATTCGTGTCAGTCAAGTTTA 1459
QY 1381 TTTCACCCAGAAAGCTCATTTTGTCTTCAACATTGTGGAATTTCTGTACTGACAC 1440
Db 1460 TTTCACCCAGAAAGCTCATTTTGTCTTCAACATTGTGGAATTTCTGTACTGACAC 1519
QY 1441 ACAGAGTTCTGCGGAGGAGATATGTAAGGCTGGCTGCTGTGTGATGCTTACGTT 1500
Db 1520 ACAGAGTTCTGCGGAGGAGATATGTAAGGCTGGCTGCTGTGTGATGCTTACGTT 1579
QY 1501 CTTGAGCCAAACATACATGATCCCATGAAAGCAGGCGAAAGCCCTGCTCTTAAGATA 1560
Db 1580 CTTGAGCCAAACATACATGATCCCATGAAAGCAGGCGAAAGCCCTGCTCTTAAGATA 1639
QY 1561 TCCATCTCTCTGAAACAATTTCTTCACTTACAGATGACCCCTCATCCCATC 1620
Db 1640 TCCATCTCTCTGAAACAATTTCTTCACTTACAGATGACCCCTCATCCCATC 1699
QY 1621 ATATATGTGTGTGCGAGAACCGGCAATGCGCGTTATGTGGTCTCTAACAATAGAG 1680
Db 1700 ATATATGTGTGTGCGAGAACCGGCAATGCGCGTTATGTGGTCTCTAACAATAGAG 1759
QY 1681 AAATCTCAAGAAACAACCCAGATGGAATTTTGAAGCATGTGTTTGTGCTGC 1740
Db 1760 AAATCTCAAGAAACAACCCAGATGGAATTTTGAAGCATGTGTTTGTGCTGC 1819
QY 1741 AGGCAATAGGATAGGATTAATCTATTACAGAAAGAGCTCAGACATTTCTTAAGATGG 1800
Db 1820 AGGCAATAGGATAGGATTAATCTATTACAGAAAGAGCTCAGACATTTCTTAAGATGG 1879
QY 1801 ATCTTAATCTATTAAGGTTTCTTCTCAAGAGATGCTCTGTGGGAGAGAGAGCC 1860
Db 1880 ATCTTAATCTATTAAGGTTTCTTCTCAAGAGATGCTCTGTGGGAGAGAGAGCC 1939
QY 1861 CCAGCAAGATATGTACAAAGACATCCAGTTTCAATGSCAGAGGTGCGAATCTCTC 1920
Db 1940 CCAGCAAGATATGTACAAAGACATCCAGTTTCAATGSCAGAGGTGCGAATCTCTC 1999
QY 1921 CTCACAGAGAGCGCCATATTTATGTGTGAGATGCAAGAAATATGSCCAAGATGTA 1980
Db 2000 CTCACAGAGAGCGCCATATTTATGTGTGAGATGCAAGAAATATGSCCAAGATGTA 2059
QY 1981 CATGATGCCCTTGTGCAATTAATAGCAAGAGTTGAGTTGAAATCTTAAGAGCAATG 2040
Db 2060 CATGATGCCCTTGTGCAATTAATAGCAAGAGTTGAGTTGAAATCTTAAGAGCAATG 2119
QY 2041 AAAACCTGGCCATTTAAAGAAAGAAACCTACTCTTCAAGATATTTGTCTATTA 2097
Db 2120 AAAACCTGGCCATTTAAAGAAAGAAACCTACTCTTCAAGATATTTGTCTATTA 2176
```

RESULT 4  
ADM43206

AD43206 standard, cDNA; 3259 BP.  
AD43206;  
03-JUN-2004 (first entry)  
Human full length cDNA encoding methionine synthase reductase.  
Human; ss; gene; Methionine synthase reductase polypeptide; HmTRR;  
cancer; cardiovascular disease; neural tube defect;  
hyperhomocysteinemia; chromosome 5p15.2-p15.3; SNP;  
single nucleotide polymorphism.  
Homo sapiens.  
Key Location/Qualifiers  
CDS 80..2176  
/\*tag= a  
/product= "hmTRR"  
variation  
/\*tag= b  
/standard\_name= "Single\_nucleotide\_polymorphism"  
variation  
/\*tag= c  
/standard\_name= "Single\_nucleotide\_polymorphism"  
US2003082676-A1.  
01-MAY-2003.  
10-AUG-1999; 99US-00371347.  
16-JAN-1998; 98US-0071622P.  
15-JAN-1999; 99US-00232028.  
(GRAV/) GRAVEL R A.  
(ROZE/) ROZEN R.  
(LECL/) LECIERC D.  
(WILS/) WILSON A.  
(ROSE/) ROSENBLATT D.  
Gravel RA, Rozen R, Lecierc D, Wilson A, Rosenblatt D;  
WPI; 2003-576610/54.  
P-PSDB; AD43207.  
New substantially pure nucleic acid encoding a mammalian methionine  
synthase reductase polypeptide, useful for diagnosing, preventing or  
treating conditions associated with altered methionine synthase activity,  
e.g. cancer.  
Example 2; SEQ ID NO 24; 26bp; English.  
The invention relates to a substantially pure nucleic acid that encodes a  
mammalian methionine synthase reductase polypeptide, HmTRR, or that  
hybridizes at high stringency to a nucleic acid appearing as AD43208 or  
AD43209. Also included are a non-human animal where one or both genetic  
alleles encoding the methionine synthase reductase polypeptide are  
mutated, an antibody that specifically binds the above methionine  
synthase reductase polypeptide, a method of detecting the presence of the  
above polypeptide, a method for detecting sequence variants for  
methionine synthase reductase in a mammal, methods of treating or  
preventing cancer (or cardiovascular disease or neural tube defects) in a  
subject, methods of screening for a compound that modulates methionine  
synthase reductase biological activity and a method for detecting an  
increased risk of developing a neural tube defect in a mammalian embryo  
or foetus. The nucleic acid is useful in diagnosing, preventing or  
treating conditions associated with altered methionine synthase activity,  
such as cancer, cardiovascular disease or neural tube defects, or in  
screening for a compound that modulates methionine synthase reductase  
biological activity. Naturally occurring variants of the polypeptide are  
also associated with hyperhomocysteinemia. The gene for HmTRR is  
located on chromosome 5p15.2-p15.3. The present sequence is full length

CC sequence of the wild-type human hmTRR cDNA.  
XX  
SQ Sequence 3259 BP; 944 A; 695 C; 674 G; 946 T; 0 U; 0 Other;  
Query Match 97.6%; Score 2046; DB 11; Length 3259;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2096; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ATGAGAGGTTCTCTTACTATATATGCTACACAGCAGGACAGGCAAGCCATCGAGAA 60  
DB 80 ATGAGAGGTTCTCTTACTATATATGCTACACAGCAGGACAGGCAAGCCATCGAGAA 139  
QY 61 GAAATGTGTGAGCAAGCTGTGTACATGATTTTTCGAGATCTTCAATATATATGAA 120  
DB 140 GAAATGTGTGAGCAAGCTGTGTACATGATTTTTCGAGATCTTCAATATATGAA 199  
QY 121 TCCGATTAAGTATGACCTTAATAAACCGAAACAGCTCCTCTGTGTGTGTTTCTACACG 180  
DB 200 TCCGATTAAGTATGACCTTAATAAACCGAAACAGCTCCTCTGTGTGTGTTTCTACACG 259  
QY 181 GGCACCGAGACCCACCCGACACAGCCCGCAAGTTTGTTAAGAAATACAGAACCAACA 240  
DB 260 GGCACCGAGACCCACCCGACACAGCCCGCAAGTTTGTTAAGAAATACAGAACCAACA 319  
QY 241 CTGCCGGTGAATTTCTTTGCTCACTCGCGTATGSGTTACTGSGTCTCGGTATTCAGAA 300  
DB 320 CTGCCGGTGAATTTCTTTGCTCACTCGCGTATGSGTTACTGSGTCTCGGTATTCAGAA 379  
QY 301 TACACCTACTTTTGGCAATGGGGGAGATTAATGATTAACGACTTCAGAGCTTGAGGCC 360  
DB 380 TACACCTACTTTTGGCAATGGGGGAGATTAATGATTAACGACTTCAGAGCTTGAGGCC 439  
QY 361 CGCATTCTTATGACACTGGACATGACATGACTGTGTAGAACTTGTGTTGAG 420  
DB 440 CGCATTCTTATGACACTGGACATGACATGACTGTGTAGAACTTGTGTTGAG 499  
QY 421 CGGTGATTTCTGATCTTGGCCAGCCCTCAGAAAGCATTTTNGTTCAGACAGACAA 480  
DB 500 CGGTGATTTCTGATCTTGGCCAGCCCTCAGAAAGCATTTTNGTTCAGACAGACAA 559  
QY 481 GAGAGATTAAGTGGGCACTCCCGTGGCATCCTGCACTCTTGAAGACAGACCTTGTG 540  
DB 560 GAGAGATTAAGTGGGCACTCCCGTGGCATCCTGCACTCTTGAAGACAGACCTTGTG 619  
QY 541 AAGTCAGAGCTGTACACATTTGATCTCAAGTCGACCTTCTGAGATTGATTCAGAGA 600  
DB 620 AAGTCAGAGCTGTACACATTTGATCTCAAGTCGACCTTCTGAGATTGATTCAGAGA 679  
QY 601 AGAAAGATTTCTGAGTTTGAAGCAAAATGCAGTGAACAGCAACCAATCCAAATGTTGA 660  
DB 680 AGAAAGATTTCTGAGTTTGAAGCAAAATGCAGTGAACAGCAACCAATCCAAATGTTGA 739  
QY 661 ATGAAGACTTTGAGTCTCACTTACCCGTTCCGTAACCCCACTTCACAAAGCCTCTGTG 720  
DB 740 ATGAAGACTTTGAGTCTCACTTACCCGTTCCGTAACCCCACTTCACAAAGCCTCTGTG 799  
QY 721 AATATTCCTGTTTACCCCAAGATATTTACAGGTACATCTGACAGAGTCTCTTGCCAG 780  
DB 800 AATATTCCTGTTTACCCCAAGATATTTACAGGTACATCTGACAGAGTCTCTTGCCAG 859  
QY 781 GAGAAAGCCCAATATCTGTGACTTCAGCAGATCCAGTTTTCAGAGTCCAAATTTCAAG 840  
DB 860 GAGAAAGCCCAATATCTGTGACTTCAGCAGATCCAGTTTTCAGAGTCCAAATTTCAAG 919  
QY 841 GCAGTTCAACTTATCTACGATGATGTCATTAATAACCACTCTGCTGATGAATTGGACAT 900  
DB 920 GCAGTTCAACTTATCTACGATGATGTCATTAATAACCACTCTGCTGATGAATTGGACAT 979  
QY 901 TCAATATACAGACTTTTCTATACAGCTGTGAGATGCTTCAGCGTGAATCTGCCTAACAGT 960  
DB 980 TCAATATACAGACTTTTCTATACAGCTGTGAGATGCTTCAGCGTGAATCTGCCTAACAGT 1039  
QY 961 GATTCTGAGGTACAAAGCCTTCTCAAAAGCTGACGTTGAAGATTAAGAGACACTGC 1020

DB 1040 GATTCTGAGGTACAAAGCCTTACCCAAAGCTGAGCTTGAGATTAAGAGAGCACTGC 1099  
1021 GTCTTTGAAAAATAAAGGACACAAAGAAAGAAAGACTTACTTACCCGACATATA 1080  
1100 GTCTTTGAAAAATAAAGGACACAAAGAAAGAAAGACTTACTTACCCGACATATA 1159  
1081 CCGGGGGAGTGTCTCTGAGTTCACTTTTACTGCTGTGAAATCCGAGCAATTCCT 1140  
1160 CCGGGGGAGTGTCTCTGAGTTCACTTTTACTGCTGTGAAATCCGAGCAATTCCT 1219  
1141 AAAAAGGCAATTTTGGAGCCCTTGTGACTATACAGTGAAGTCTGAAAAAGCCAG 1200  
1220 AAAAAGGCAATTTTGGAGCCCTTGTGACTATACAGTGAAGTCTGAAAAAGCCAG 1279  
1201 CTACAGAGCTGTGACGTAAACAGAGGGGACCCGATTAATAGCCGCTTTGTACGAGTCC 1260  
1280 CTACAGAGCTGTGACGTAAACAGAGGGGACCCGATTAATAGCCGCTTTGTACGAGTCC 1339  
1261 TGTGCTGTGTGTGATCTCTCTCTGCTTCCCTTCCCTTCCGAGCACCACCTCACTTC 1320  
1340 TGTGCTGTGTGTGATCTCTCTCTGCTTCCCTTCCCTTCCGAGCACCACCTCACTTC 1399  
1321 CTGCTGAAATCTTCTTAACTTCAACCCAGACATATTCGTGTGCAAGCTCAAGTTTA 1380  
1400 CTGCTGAAATCTTCTTAACTTCAACCCAGACATATTCGTGTGCAAGCTCAAGTTTA 1459  
1381 TTTCACCCAGAAAGCTCAATTTTGTCTTCAACATTTGTGAAATTTCTGTCTACGCA 1440  
1460 TTTCACCCAGAAAGCTCAATTTTGTCTTCAACATTTGTGAAATTTCTGTCTACGCA 1519  
1441 ACAAGAGTTCTGGGAGAGAGATATGATACAGCTGTGCTGCTTGTGTGTGCTTCA 1500  
1520 ACAAGAGTTCTGGGAGAGAGATATGATACAGCTGTGCTGCTTGTGTGTGCTTCA 1579  
1501 CTTGAGCCAAACATACATGATCCCATGAGAGACAGGGGAAAGCCCTGAGCTCTTAAGATA 1560  
1580 CTTGAGCCAAACATACATGATCCCATGAGAGACAGGGGAAAGCCCTGAGCTCTTAAGATA 1639  
1561 TCCATCTCTCTCTGAAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
1640 TCCATCTCTCTCTGAAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1699  
1621 ATATATGTGGGTCTCAGAAACCCGACATAGCCCGTTTATTTGGTCTCTAACAATAGAG 1680  
1700 ATATATGTGGGTCTCAGAAACCCGACATAGCCCGTTTATTTGGTCTCTAACAATAGAG 1759  
1681 AAACCTCAAGAAACAACCCAGATGAGAAATTTGAGCAATGTGTGTTTTTGGCTGC 1740  
1760 AAACCTCAAGAAACAACCCAGATGAGAAATTTGAGCAATGTGTGTTTTTGGCTGC 1819  
1741 AGGCATAGAGATAGGATTAATCTATTCAAGAAAGACTGAGATTTCTTAAGCATGG 1800  
1820 AGGCATAGAGATAGGATTAATCTATTCAAGAAAGACTGAGATTTCTTAAGCATGG 1879  
1801 ATCTTAATCTATTAAGGTTTCTCTCAAGAGATGCTCTGTGGGAGAGAGAGCC 1860  
1880 ATCTTAATCTATTAAGGTTTCTCTCAAGAGATGCTCTGTGGGAGAGAGAGCC 1939  
1861 CCAAGCAAGATATGACAGCAACATCCAGCTTCAATGCGCAGAGAGGAGAGATCTTC 1920  
1940 CCAAGCAAGATATGACAGCAACATCCAGCTTCAATGCGCAGAGAGGAGAGATCTTC 1999  
1921 CTCACAGAGACCGCATATTTATGTGTGTGAGATGCAAGAAATATGCGCAAGATGTA 1980  
2000 CTCACAGAGACCGCATATTTATGTGTGTGAGATGCAAGAAATATGCGCAAGATGTA 2059  
1981 CATGATGCTTGTGCAATATATAGCAAGAGGTTGAGTTGAAAACTAGAGCAATG 2040  
2060 CATGATGCTTGTGCAATATATAGCAAGAGGTTGAGTTGAAAACTAGAGCAATG 2119  
2041 AAAACCTGCGCACTTTAAAGAAAGAAAGGCTACCTTCAAGATTTTGTGATATA 2097

DB 2120 AAAACCTGCGCACTTTAAAGAAAGAAAGAAAGGCTTACAGATATTTGTCAATA 2176  
RESULT 5  
ADM43208  
ID ADM43208 standard; cDNA; 2094 BP.  
XX  
AC ADM43208;  
XX  
DT 03-JUN-2004 (first entry)  
XX  
DE Human wild-type methionine synthase reductase CDS.  
XX  
KW Human; ss; Methionine synthase reductase polypeptide; HsMTRR; cancer;  
KW cardiovascular disease; neural tube defect; hyperhomocysteinemia;  
KW chromosome 5p15.2-p15.3; SNP; single nucleotide polymorphism.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..2094  
FT /tag= a  
FT /product= "HsMTRR"  
FT /partial  
FT /note= "No stop codon shown"  
FT variation  
FT /tag= b  
FT /standard\_name= "Single\_nucleotide\_polymorphism"  
FT replace(110,A)  
FT /\*tag= C  
FT /standard\_name= "Single\_nucleotide\_polymorphism"  
XX  
EN US2003082676-A1.  
XX  
PD 01-MAY-2003.  
XX  
PF 10-AUG-1999; 99US-00371347.  
XX  
PR 16-JAN-1998; 98US-00716222.  
PR 15-JAN-1999; 99US-00232028.  
XX  
PA (GRAV/) GRAVEL R A.  
PA (ROZE/) ROZEN R.  
PA (LECL/) LECLEERC D.  
PA (WILS/) WILSON A.  
PA (ROSE/) ROSENBLATT D.  
XX  
PI Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;  
XX  
DR WPI, 2003-576610/54.  
DR P-PSDB; ADM43207.  
XX  
PT New substantially pure nucleic acid encoding a mammalian methionine  
PT synthase reductase polypeptide, useful for diagnosing, preventing or  
PT treating conditions associated with altered methionine synthase activity,  
PT e.g. cancer.  
XX  
PS Claim 3; SEQ ID NO 1; 26pp; English.  
XX  
CC The invention relates to a substantially pure nucleic acid that encodes a  
CC mammalian methionine synthase reductase polypeptide, HsMTRR, or that  
CC hybridizes at high stringency to a nucleic acid appearing as ADM43208 or  
CC ADM43209. Also included are a non-human animal where one or both genetic  
CC alleles encoding the methionine synthase reductase polypeptide are  
CC mutated, an antibody that specifically binds the above methionine  
CC synthase reductase polypeptide, a method of detecting the presence of the  
CC above polypeptide, a method for detecting sequence variants for  
CC methionine synthase reductase in a mammal, methods of treating or  
CC preventing cancer (or cardiovascular disease or neural tube defects) in a  
CC subject, methods of screening for a compound that modulates methionine  
CC synthase reductase biological activity and a method for detecting an  
CC increased risk of developing a neural tube defect in a mammalian embryo  
CC or foetus. The nucleic acid is useful in diagnosing, preventing or

CC treating conditions associated with altered methionine synthase activity,  
CC such as cancer, cardiovascular disease or neural tube defects, or in  
CC screening for a compound that modulates methionine synthase reductase  
CC biological activity. Naturally occurring variants of the polypeptide are  
CC also associated with hyperhomocysteinaemia. The gene for HMTMR is  
CC located on chromosome 5p15.2-p15.3. The present sequence is the coding  
CC sequence of the wild-type human hMTMR cDNA.

XX Sequence 2094 BP; 591 A; 489 C; 481 G; 533 T; 0 U; 0 Other;

Query Match 97.4%; Score 2043; DB 11; Length 2094;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGAGAGGTTTCTGTACTATATGCTACACAGGAGGACAGGCAAGCCATCGAGAA 60  
1 ATGAGAGGTTTCTGTACTATATGCTACACAGGAGGACAGGCAAGCCATCGAGAA 60  
61 GAAATGTGTAGCAAGCTGTGTATCATGATTTTCTGCAATCTTCACTATATTAAGAA 120  
61 GAAATGTGTAGCAAGCTGTGTATCATGATTTTCTGCAATCTTCACTATATTAAGAA 120  
121 TCCGATTAAGTATGACTTAAACCGGAAACAGCTCTCTGTTGTTGTGTTTCTACACAG 180  
121 TCCGATTAAGTATGACTTAAACCGGAAACAGCTCTCTGTTGTTGTGTTTCTACACAG 180  
181 GGCACCGGAGACCCACCCGACACAGCCCGCAAGTTTGTAAAGAAATACAAACAAACA 240  
181 GGCACCGGAGACCCACCCGACACAGCCCGCAAGTTTGTAAAGAAATACAAACAAACA 240  
241 CTGCGGTTTATTTCTTGTCTACACCTGCGATAGGTTTACTGGGTCTCGGTGATTCGAA 300  
241 CTGCGGTTTATTTCTTGTCTACACCTGCGATAGGTTTACTGGGTCTCGGTGATTCGAA 300  
301 TACACTACTTTTTCGAATGGGGGAGAGATTAATGATAACGACTTCAAGAGCTTGAAGCC 360  
301 TACACTACTTTTTCGAATGGGGGAGAGATTAATGATAACGACTTCAAGAGCTTGAAGCC 360  
361 CGGCAATTTATGACACTGACATGACATGACATGCTGTAGGTTTGAATCTTGTTGAG 420  
361 CGGCAATTTATGACACTGACATGACATGACATGCTGTAGGTTTGAATCTTGTTGAG 420  
421 CCGTGATTTGCTGACCTCTGCGCAGCCCTCAGAAAGCATTTTATGATCAAGCAGAGCAA 480  
421 CCGTGATTTGCTGACCTCTGCGCAGCCCTCAGAAAGCATTTTATGATCAAGCAGAGCAA 480  
481 GAGGAGTAAGTGTGCGACCTCCGGTGGACATCACTGATCTTGAAGAGCAGCTTGTG 540  
481 GAGGAGTAAGTGTGCGACCTCCGGTGGACATCACTGATCTTGAAGAGCAGCTTGTG 540  
541 AAGTCAGAGCTGTACATTAATTAATTCAGAGTCTTGAAGTTGATGATTCAGGA 600  
541 AAGTCAGAGCTGTACATTAATTAATTCAGAGTCTTGAAGTTGATGATTCAGGA 600  
601 AGAAGAGATTTCTGAGTTTGAAGCAAAATGCAATGAAAGCAACCAATCCAAATGTTGA 660  
601 AGAAGAGATTTCTGAGTTTGAAGCAAAATGCAATGAAAGCAACCAATCCAAATGTTGA 660  
661 ATTGAAGATTTGAGTCTCACTTACCCTTCCGTTCCGTTACCCCACTCTCAAGAGCTCTG 720  
661 ATTGAAGATTTGAGTCTCACTTACCCTTCCGTTCCGTTACCCCACTCTCAAGAGCTCTG 720  
721 AATATCTCTGTTTACCCCGAATATTTTACAGGTATCTGACGAGAGTCTCTTGAGCAG 780  
721 AATATCTCTGTTTACCCCGAATATTTTACAGGTATCTGACGAGAGTCTCTTGAGCAG 780  
781 GAGGAAGCCCAAGATCTGTGACATTTCAAGATTCAGATTTTCAAGTCCAAATTTCAAG 840  
781 GAGGAAGCCCAAGATCTGTGACATTTCAAGATTCAGATTTTCAAGTCCAAATTTCAAG 840  
841 GCAGTTCACTTACAGATGATGACATAAAACCACTGCTGTGTGATTTGAGATTTGAGATT 900  
841 GCAGTTCACTTACAGATGATGACATAAAACCACTGCTGTGTGATTTGAGATTTGAGATT 900

901 TCAATACAGACTTTTCTATACAGCTGTGAGATGCTTCAAGCTGATCTGCCATACAGT 960  
901 TCAATACAGACTTTTCTATACAGCTGTGAGATGCTTCAAGCTGATCTGCCATACAGT 960  
961 GATTCGAGGTCAAGGCTTACCCCAAGATCTGACCTTGAAGATTAAGAGAGCAGCTGC 1020  
961 GATTCGAGGTCAAGGCTTACCCCAAGATCTGACCTTGAAGATTAAGAGAGCAGCTGC 1020  
1021 GTCCCTTTGAAATTAAGGACAGACACAAAGAAAGAGCTTACCTTACCCAGATATA 1080  
1021 GTCCCTTTGAAATTAAGGACAGACACAAAGAAAGAGCTTACCTTACCCAGATATA 1080  
1081 CTGCGGAGATGTTCTCTCAAGTTCATTTTACCTGTTCTTGAATTCGAGCAATTTCTT 1140  
1081 CTGCGGAGATGTTCTCTCAAGTTCATTTTACCTGTTCTTGAATTCGAGCAATTTCTT 1140  
1141 AAAAAGCAATTTTGGAGCCCTGTGGACTATACAGTACAGTCTGAAAAAGCCAGG 1200  
1141 AAAAAGCAATTTTGGAGCCCTGTGGACTATACAGTACAGTCTGAAAAAGCCAGG 1200  
1201 CTACAGAGCTGTGACATTAACAAAGGAGGAGCCGATTAAGCCGTTTGTACAGATGCC 1260  
1201 CTACAGAGCTGTGACATTAACAAAGGAGGAGCCGATTAAGCCGTTTGTACAGATGCC 1260  
1261 TGTGCTGCTTGTGATCTCTCTGCTGCTTCTTCTTGTGAGGACCACTCACTGCTC 1320  
1261 TGTGCTGCTTGTGATCTCTCTGCTGCTTCTTCTTGTGAGGACCACTCACTGCTC 1320  
1321 CTGCGAATCTTCTTAACTTCAACCCAGACATATTCGTTGCAAGCTCAAGTTTA 1380  
1321 CTGCGAATCTTCTTAACTTCAACCCAGACATATTCGTTGCAAGCTCAAGTTTA 1380  
1381 TTTCAACCAAGAAAGCTCAATTTTGTCTTCAATTTGTGATTTGTCTACTGACACA 1440  
1381 TTTCAACCAAGAAAGCTCAATTTTGTCTTCAATTTGTGATTTGTCTACTGACACA 1440  
1441 ACAAGGTTCTGCGAAGGAGATGATACAGCTGCTGCTGCTTGTGTTGCTTCACT 1500  
1441 ACAAGGTTCTGCGAAGGAGATGATACAGCTGCTGCTGCTTGTGTTGCTTCACT 1500  
1501 CTTCAAGCAAACTATATGATGATCCATGAAAGACAGGAGGAAAGCCCTGCTCTTAAGATA 1560  
1501 CTTCAAGCAAACTATATGATGATCCATGAAAGACAGGAGGAAAGCCCTGCTCTTAAGATA 1560  
1561 TCCATCTCTCTGAAACCAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
1561 TCCATCTCTCTGAAACCAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
1621 ATATAGTGGGTCAAGAACCCGATAGCCCGTTTATTTGGTTCTTCAACATTAAGAG 1680  
1621 ATATAGTGGGTCAAGAACCCGATAGCCCGTTTATTTGGTTCTTCAACATTAAGAG 1680  
1681 AAACTCAAGAAACAACCCAGATGAAATTTTGGACAAATGTTGTTTGTGCTG 1740  
1681 AAACTCAAGAAACAACCCAGATGAAATTTTGGACAAATGTTGTTTGTGCTG 1740  
1741 AGCATAAGAGTAAGGATTAATCTATTCAGAAAAAGCTCAGACATTTCTTAAAGATGG 1800  
1741 AGCATAAGAGTAAGGATTAATCTATTCAGAAAAAGCTCAGACATTTCTTAAAGATGG 1800  
1801 AGGATTAAGGATTAAGGATTAATCTATTCAGAAAAAGCTCAGACATTTCTTAAAGATGG 1860  
1801 AGGATTAAGGATTAAGGATTAATCTATTCAGAAAAAGCTCAGACATTTCTTAAAGATGG 1860  
1861 CCAAGCAAGATTAAGCAACATCCAGCTTCAAGCCAGAGGAGGAGGAGATCTC 1920  
1861 CCAAGCAAGATTAAGCAACATCCAGCTTCAAGCCAGAGGAGGAGGAGATCTC 1920  
1921 CTCAGAGAAAGCGCATATTTATGTTGTGAGATGCAAGAAATATGCGCAAGATGTA 1980  
1921 CTCAGAGAAAGCGCATATTTATGTTGTGAGATGCAAGAAATATGCGCAAGATGTA 1980

QY 1981 CATGATGCTCTTGTGCAATATATAGCAAGAGTTGAGTTGAAAACTAGAGCAATG 2040  
 DB 1981 CATGATGCTCTTGTGCAATATATAGCAAGAGTTGAGTTGAAAACTAGAGCAATG 2040  
 QY 2041 AAAACCTGGCCACTTTAAAGAAAGAAAGCGTACCTTCAGATATTTGGTCA 2094  
 DB 2041 AAAACCTGGCCACTTTAAAGAAAGAAAGCGTACCTTCAGATATTTGGTCA 2094  
 RESULT 6  
 ADM43209  
 ID ADM43209 standard; cDNA; 2094 BP.  
 AC ADM43209;  
 XX  
 DT 03-JUN-2004 (first entry)  
 XX  
 DE Human methionine synthase reductase CDS G66A variant.  
 XX  
 KM Human; ss; Methionine synthase reductase polypeptide; HsmTRR; cancer;  
 XX cardiovascular disease; neural tube defect; hyperhomocysteinemia;  
 KM chromosome 5p15.2-p15.3; SNP; single nucleotide polymorphism.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FH CDS 1..2094  
 FT /\*tag= a  
 FT /product= "hsmTRR"  
 FT /partial  
 FT /note= "No stop codon shown"  
 FT replace(66, G)  
 FT /\*tag= b  
 FT /standard\_name= "Single\_nucleotide\_polymorphism"  
 FT /replace(110, A)  
 FT /\*tag= c  
 FT /standard\_name= "Single\_nucleotide\_polymorphism"  
 FT US2003082676-A1.  
 XX  
 PN 01-MAY-2003.  
 PD 10-AUG-1999; 99US-00371347.  
 PF 16-JAN-1998; 98US-0071622P.  
 PR 15-JAN-1999; 99US-00232028.  
 XX  
 PA (GRAV/) GRAVEL R A.  
 PA (ROZE/) ROZEN R.  
 PA (LECL/) LECLERC D.  
 PA (WILS/) WILSON A.  
 PA (ROSE/) ROSENBLATT D.  
 PI Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;  
 XX  
 DR WPI; 2003-576610/54.  
 DR P-PSDB; ADM43211.  
 XX  
 PT New substantially pure nucleic acid encoding a mammalian methionine  
 PT synthase reductase polypeptide, useful for diagnosing, preventing or  
 PT treating conditions associated with altered methionine synthase activity,  
 PT e.g. cancer.  
 XX  
 PS Claim 3; SEQ ID NO 41; 26bp; English.  
 XX  
 CC The invention relates to a substantially pure nucleic acid that encodes a  
 CC mammalian methionine synthase reductase polypeptide. HsmTRR, or that  
 CC hybridizes at high stringency to a nucleic acid appearing as ADM43208 or  
 CC ADM43209. Also included are a non-human animal where one or both genetic  
 CC alleles encoding the methionine synthase reductase polypeptide are  
 CC mutated, an antibody that specifically binds the above methionine  
 CC synthase reductase polypeptide, a method of detecting the presence of the  
 CC above polypeptide, a method for detecting sequence variants for

CC methionine synthase reductase in a mammal, methods of treating or  
 CC preventing cancer (or cardiovascular disease or neural tube defects) in a  
 CC subject, methods of screening for a compound that modulates methionine  
 CC synthase reductase biological activity and a method for detecting an  
 CC increased risk of developing a neural tube defect in a mammalian embryo  
 CC or foetus. The nucleic acid is useful in diagnosing, preventing or  
 CC treating conditions associated with altered methionine synthase activity,  
 CC such as cancer, cardiovascular disease or neural tube defects, or in  
 CC screening for a compound that modulates methionine synthase reductase  
 CC biological activity. Naturally occurring variants of the polypeptide are  
 CC also associated with hyperhomocysteinemia. The gene for HsmTRR is  
 CC located on chromosome 5p15.2-p15.3. The present sequence is the coding  
 CC sequence of a variant human hsmTRR cDNA.  
 XX  
 SQ Sequence 2094 BP; 592 A; 489 C; 480 G; 533 T; 0 U; 0 Other;  
 Query Match 95.0%; Score 1992; DB 11; Length 2094;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 2092; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 ATGAGAGAGTTTCTGTACTATATATGCTACACAGCAGGAGCAAGGCAATCGCAGAA 60  
 DB 1 ATGAGAGAGTTTCTGTACTATATATGCTACACAGCAGGAGCAAGGCAATCGCAGAA 60  
 QY 61 GAAATGTGAGCAAGCTGTGTAATGATTTTCTGAGATCTTCACTATATAGTGA 120  
 DB 61 GAAATGTGAGCAAGCTGTGTAATGATTTTCTGAGATCTTCACTATATAGTGA 120  
 QY 121 TCCGATATGATGACCTAATAAACCAGAACAGCTCTCTTGTGTGTGTTTCTACACG 180  
 DB 121 TCCGATATGATGACCTAATAAACCAGAACAGCTCTCTTGTGTGTGTTTCTACACG 180  
 QY 181 GGACCCGAGACCCACCCGACACAGCCCGCAAGTTTGAAGAAATACAGAACCAAC 240  
 DB 181 GGACCCGAGACCCACCCGACACAGCCCGCAAGTTTGAAGAAATACAGAACCAAC 240  
 QY 181 GGACCCGAGACCCACCCGACACAGCCCGCAAGTTTGAAGAAATACAGAACCAAC 240  
 DB 181 GGACCCGAGACCCACCCGACACAGCCCGCAAGTTTGAAGAAATACAGAACCAAC 240  
 QY 241 CTGCGGTGATTTCTTCTGCTACCTGCGGTATGAGTTTCTGAGTCTCGGTATTCAGAA 300  
 DB 241 CTGCGGTGATTTCTTCTGCTACCTGCGGTATGAGTTTCTGAGTCTCGGTATTCAGAA 300  
 QY 301 TACACCTACTTTTGCAATGGGGGAAAGATATTTGATTAACACTTCAAGAGCTTGAGCC 360  
 DB 301 TACACCTACTTTTGCAATGGGGGAAAGATATTTGATTAACACTTCAAGAGCTTGAGCC 360  
 QY 361 CGGCAATTTCTATGACACTGGAATGAGTACTGTGAGTTTGAATCTTGCGTTGAG 420  
 DB 361 CGGCAATTTCTATGACACTGGAATGAGTACTGTGAGTTTGAATCTTGCGTTGAG 420  
 QY 421 CCGTGAATGCTGGACTCTGCGCAGCCCTCAGAAAGCATTTTATAGTCAAGAGAGACAA 480  
 DB 421 CCGTGAATGCTGGACTCTGCGCAGCCCTCAGAAAGCATTTTATAGTCAAGAGAGACAA 480  
 QY 481 GAGGAGATAGTGGCGCACTCCCGGTGATCACTGCAATCTTGAAGACAGACTTGTG 540  
 DB 481 GAGGAGATAGTGGCGCACTCCCGGTGATCACTGCAATCTTGAAGACAGACTTGTG 540  
 QY 541 AAGTCAGAGCTGTACACATTTGATTTCAAGTGTGAGCTTCTGAGATTCATGATTCAGAA 600  
 DB 541 AAGTCAGAGCTGTACACATTTGATTTCAAGTGTGAGCTTCTGAGATTCATGATTCAGAA 600  
 QY 601 AGAAGAGTTTGAAGTTTGAAGCAAAATGAGTGAACAGCAACCAATCCAAATGTTGTA 660  
 DB 601 AGAAGAGTTTGAAGTTTGAAGCAAAATGAGTGAACAGCAACCAATCCAAATGTTGTA 660  
 QY 661 ATTGAAGACTTGAGTCTCACTTACCCGTTGATTCAGAGCTCTGAGAGCTTCTG 720  
 DB 661 ATTGAAGACTTGAGTCTCACTTACCCGTTGATTCAGAGCTCTGAGAGCTTCTG 720  
 QY 721 AATATTCCTGTTTACCCCAAGAAATTTATACAGTCACTGAGAGAGTCTTGCGCCAG 780  
 DB 721 AATATTCCTGTTTACCCCAAGAAATTTATACAGTCACTGAGAGAGTCTTGCGCCAG 780  
 QY 781 GAGGAAGCAAGATATCTGTGACTTCAGACAGATCCAGTTTTCAGAGTGCATTTCAAG 840

```
Db 781 GAGAGAAAGCCAGATATCTGACTGACGACATCCATTTTTCAGTGCCAAATTTCAAAG 840
Qy 841 GCAGTTCAACTTACTAGAAATGATGCCATTAACCACTGCTGTAGATTTGACATT 900
Db 841 GCAGTTCAACTTACTAGAAATGATGCCATTAACCACTGCTGTAGATTTGACATT 900
Qy 901 TCMAATACAGACTTTTCTTATCAGCTTGAGATGCTTCAAGCTGATCTCCCTAACAT 960
Db 901 TCMAATACAGACTTTTCTTATCAGCTTGAGATGCTTCAAGCTGATCTCCCTAACAT 960
Qy 961 GATTCTGAGGTACAAAGCCTACTCCAAAGA CTGACGCTTGAAGATAAAGAGACCTGC 1020
Db 961 GATTCTGAGGTACAAAGCCTACTCCAAAGA CTGACGCTTGAAGATAAAGAGACCTGC 1020
Qy 1021 GTCTTTTGAATAATTAAGGAGACACAAAGAAAGAGAGACTTACCTTACCCAGCATATA 1080
Db 1021 GTCTTTTGAATAATTAAGGAGACACAAAGAAAGAGAGACTTACCTTACCCAGCATATA 1080
Qy 1081 CTGCGGGAATGTTCTCTCAGTTCATTTTAACTGTGTCTTGAATCCGACAAATTCCT 1140
Db 1081 CTGCGGGAATGTTCTCTCAGTTCATTTTAACTGTGTCTTGAATCCGACAAATTCCT 1140
Qy 1141 AAAAAGGCAATTTTGGAGCCCTGTGACATTAACAGTGAAGAGTGAAGAGCCGAGG 1200
Db 1141 AAAAAGGCAATTTTGGAGCCCTGTGACATTAACAGTGAAGAGTGAAGAGCCGAGG 1200
Qy 1201 CTACAGAGCTGTGCACTTAAACAAAGGGGAGCCGATTAATAGCCGCTTGTACGAGATGCC 1260
Db 1201 CTACAGAGCTGTGCACTTAAACAAAGGGGAGCCGATTAATAGCCGCTTGTACGAGATGCC 1260
Qy 1261 TGTGCTGCTGTTGTGATCTCTCTCTGCTTCCCTTCTTGCCAGCCACATCACTCTC 1320
Db 1261 TGTGCTGCTGTTGTGATCTCTCTCTGCTTCCCTTCTTGCCAGCCACATCACTCTC 1320
Qy 1321 CTGCTCGAATCTTCTCTTAACTTCAACCCAGACATTAATGATGAGAGTCAAGTTTA 1380
Db 1321 CTGCTCGAATCTTCTCTTAACTTCAACCCAGACATTAATGATGAGAGTCAAGTTTA 1380
Qy 1381 TTTCACCCAGAAAGCTCAATTTTGTCTCAACATTTGTGATTTCTGTCTACTGCCACA 1440
Db 1381 TTTCACCCAGAAAGCTCAATTTTGTCTCAACATTTGTGATTTCTGTCTACTGCCACA 1440
Qy 1441 ACAGAGTTCGCGGAAAGGAGTATGTAACAGCTGCGCTGCTGTTGTTGCTTCACTT 1500
Db 1441 ACAGAGTTCGCGGAAAGGAGTATGTAACAGCTGCGCTGCTGTTGTTGCTTCACTT 1500
Qy 1501 CTTGAGCCAAACATACATGATGCCATGAAGAAGCGGAAAGCCCTGCTCCCTAAGATA 1560
Db 1501 CTTGAGCCAAACATACATGATGCCATGAAGAAGCGGAAAGCCCTGCTCCCTAAGATA 1560
Qy 1561 TCCATCTCTCTCGAACAACAATTTCTTCCACTTACAGATGAACCCCTCAATCCCATC 1620
Db 1561 TCCATCTCTCTCGAACAACAATTTCTTCCACTTACAGATGAACCCCTCAATCCCATC 1620
Qy 1621 ATTAATGTTGGTCCAGAGACCGGATAGCCCCGTTTATGAGGTTCTTAACAATAGAG 1680
Db 1621 ATTAATGTTGGTCCAGAGACCGGATAGCCCCGTTTATGAGGTTCTTAACAATAGAG 1680
Qy 1681 AAAATCCAAAGAAACACCCAGATGAAATTTTGAGCAATGAGTGTGTTTGGCTGC 1740
Db 1681 AAAATCCAAAGAAACACCCAGATGAAATTTTGAGCAATGAGTGTGTTTGGCTGC 1740
Qy 1741 AGGCAATAGATAGGATATATCTATTCAGAAAGAGCTCAGACATTTCTTAAAGCATGG 1800
Db 1741 AGGCAATAGATAGGATATATCTATTCAGAAAGAGCTCAGACATTTCTTAAAGCATGG 1800
Qy 1801 ATCTTAATCATCTTAAAGGTTTCTTCTCAAGAGATGCTCTGTTGGGAGAGAGAGCC 1860
Db 1801 ATCTTAATCATCTTAAAGGTTTCTTCTCAAGAGATGCTCTGTTGGGAGAGAGAGCC 1860
Qy 1861 CCAGCAAGTATATTAAGAGACATCCAGCTTCAATGCGCAGAGTGGCGAGAAATCTCTC 1920
Db 1861 CCAGCAAGTATATTAAGAGACATCCAGCTTCAATGCGCAGAGTGGCGAGAAATCTCTC 1920
```

```
Db 1861 CCAGCAAGTATATTAAGAGACATCCAGCTTCAATGCGCAGAGTGGCGAGATCTCTC 1920
Qy 1921 CTCAGAGAGAGCGCCATATTTATGTGTGAGATGCAAGAAATATGCGCAAGATGTA 1980
Db 1921 CTCAGAGAGAGCGCCATATTTATGTGTGAGATGCAAGAAATATGCGCAAGATGTA 1980
Qy 1981 CATGATGCCCTTGTGAAATTAATTAACCAAGAGGTTGAGTTGAAAACCTAGAGCAATG 2040
Db 1981 CATGATGCCCTTGTGAAATTAATTAACCAAGAGGTTGAGTTGAAAACCTAGAGCAATG 2040
Qy 2041 AAAACCTGCGCACTTTAAAGAAAGAAACGCTACCTTCAGAGTATTTGTCA 2094
Db 2041 AAAACCTGCGCACTTTAAAGAAAGAAACGCTACCTTCAGAGTATTTGTCA 2094

RESULT 7
ID AA58935
AA58935 standard; DNA; 3259 BP.
AC
XX
AC AA58935;
XX
DT 07-NOV-2000 (first entry)
XX
DE DNA encoding a human methionine synthase reductase polypeptide.
XX
KW Human; methionine synthase reductase; MTRR; cancer;
KW cardiovascular disease; Down's Syndrome; neural tube defect;
KW premature coronary artery disease; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 80..2176
FT /*tag= a
FT /product= "methionine synthase reductase"
XX
PN NC0200042196-A2.
XX
PD 20-JUL-2000.
XX
PF 14-JAN-2000; 2000MO-IB000209.
XX
PR 15-JAN-1999; 99US-00232028.
PR 10-AUG-1999; 99US-00371347.
XX
PA (UYMC-) UNIV MCGILL.
XX
PI Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;
XX
DR WPI, 2000-46131/40.
DR P-PSDB; AAB07591.
XX
XX
PT Mammalian methionine synthase reductase nucleic acid used for detecting
PT an increased risk of developing a neural tube defect, Down's Syndrome or
PT cardiovascular disease in a mammalian embryo or fetus.
XX
PS Claim 3; Fig 3; 85bp; English.
XX
CC The present sequence encodes a human methionine synthase reductase (MTRR)
CC polypeptide. Inhibitors of MTRR polypeptide and polynucleotide are used
CC for treating or preventing cancer, cardiovascular disease, Down's
CC Syndrome or neural tube defects in a subject. The cardiovascular disease
CC is premature coronary artery disease. The compounds are detected by
CC methods which screen for modulators of MTRR biological activity. MTRR
CC polypeptide or nucleic acid is examined for the presence of a
CC polymorphism in the parents or the embryo or foetus, and the information
CC used for detecting an increased risk of an embryo or foetus developing
CC cancer, cardiovascular disease, Down's Syndrome or neural tube defects
XX
SQ Sequence 3259 BP; 944 A; 706 C; 663 G; 946 T; 0 U; 0 Other;
Query Match 92.7%; Score 1944; DB 3; Length 3259;
Best Local Similarity 99.9%; Pred. No. 0;
```



Matches 2094; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
OY 1 ATGAGGAGGTTCTGTACTATATGCTACACAGGAGGACAGGCAAAAGGCATCGAGAA 60
DB 80 ATGAGGAGGTTCTGTACTATATGCTACACAGGAGGACAGGCAAAAGGCATCGAGAA 139
OY 61 GAAATGTGAGACAGCTGTGTAATGATTTTCTGAGATCTTCACTATATTAAGTAA 120
DB 140 GAAATGTGAGACAGCTGTGTAATGATTTTCTGAGATCTTCACTATATTAAGTAA 199
OY 121 TCCGATAGTATGACCTTAAACCCGAAACAGCTCTCTTGTGTGTGTGTCTACACG 180
DB 200 TCCGATAGTATGACCTTAAACCCGAAACAGCTCTCTTGTGTGTGTGTCTACACG 259
OY 181 GGCAACCGGAGACCCACCCGACACAGCCCGCAAGTTGTTAAGAAATTAAGAAACCA 240
DB 260 GGCAACCGGAGACCCACCCGACACAGCCCGCAAGTTGTTAAGAAATTAAGAAACCA 319
OY 241 CTGCCGTTGATTTCTTCTCTCACTGCGGATGAGTTACTGGGTCTCGGTGATTGAAA 300
DB 320 CTGCCGTTGATTTCTTCTCTCACTGCGGATGAGTTACTGGGTCTCGGTGATTGAAA 379
OY 301 TACACCTACTTTTGAAGTGGGGGAAAGATTAATTGATTAACGACTTCAAGAGCTTGAGCC 360
DB 380 TACACCTACTTTTGAAGTGGGGGAAAGATTAATTGATTAACGACTTCAAGAGCTTGAGCC 439
OY 361 CGGCAATTTCTATGACCTGGAACATGCAATGATGATGATGATTAAGCTTGTGTGAG 420
DB 440 CGGCAATTTCTATGACCTGGAACATGCAATGATGATGATGATTAAGCTTGTGTGAG 499
OY 421 CCGTGAATTTGAGACTGTCGACGACCTCAGAAACATTTTATGATCAAGCAAGAGCAA 480
DB 500 CCGTGAATTTGAGACTGTCGACGACCTCAGAAACATTTTATGATCAAGCAAGAGCAA 559
OY 481 GAGGAGATAAGTGGCGACCTCCGGTGGCATCACTGCATCTTGAAGACAGACCTTGTG 540
DB 560 GAGGAGATAAGTGGCGACCTCCGGTGGCATCACTGCATCTTGAAGACAGACCTTGTG 619
OY 541 AAGTCAGAGCTGTACACATTTGAACTTCAAGTGAAGCTTGTGATTTGATGATTTCAAGA 600
DB 620 AAGTCAGAGCTGTACACATTTGAACTTCAAGTGAAGCTTGTGATTTGATGATTTCAAGA 679
OY 601 AGAAGAGATTTGAGGTTTGAAGCAAAATGACGTGAACAGCAACCAATCCATGTTGTA 660
DB 680 AGAAGAGATTTGAGGTTTGAAGCAAAATGACGTGAACAGCAACCAATCCATGTTGTA 739
OY 661 ATTGAAGACTTTGAGTCTCACTTACCCGTTCCGTAACCCCACTCTCAAGCCCTCTG 720
DB 740 ATTGAAGACTTTGAGTCTCACTTACCCGTTCCGTAACCCCACTCTCTCAAGCCCTCTG 799
OY 721 AATATTCCTGTTTACCCCAAAATTTTACAGGTACATGTCAGAGAGTCTCTTGGCCAG 780
DB 800 AATATTCCTGTTTACCCCAAAATTTTACAGGTACATGTCAGAGAGTCTCTTGGCCAG 859
OY 781 GAGGAAAGCAAGTATCTGTGACTTCAAGAGATCCAGTTTCAAGTCCAAATTTCAAG 840
DB 860 GAGGAAAGCAAGTATCTGTGACTTCAAGAGATCCAGTTTCAAGTCCAAATTTCAAG 919
OY 841 GCAGTTCAACTTACTACGATGATGCAATTAACCACTCTGCTGTAGTAATTTGACAT 900
DB 920 GCAGTTCAACTTACTACGATGATGCAATTAACCACTCTGCTGTAGTAATTTGACAT 979
OY 901 TCAATATCAGACTTTTCTCTATGAGCTGAGATGCTTCAAGCTGATCTGCTTAAACAT 960
DB 980 TCAATATCAGACTTTTCTCTATGAGCTGAGATGCTTCAAGCTGATCTGCTTAAACAT 1039
OY 961 GATTCGAGGTACAAAGCTTACTCAAAAGCTGAGCTTGAAGATTAAGAGAGCACTGC 1020
DB 1040 GATTCGAGGTACAAAGCTTACTCAAAAGCTGAGCTTGAAGATTAAGAGAGCACTGC 1099
OY 1021 GTCTCTTTGAAATTAAGGACGACACAAAGAAAGAGAGCTTACCCGACATATA 1080
DB 1100 GTCTCTTTGAAATTAAGGACGACACAAAGAAAGAGAGCTTACCCGACATATA 1159
```

```
OY 1081 CCGCGGAGATGTTCTCTCCAGTTCAATTTTACCCTGTGTCTTGAATCCGAGCAATTCCT 1140
DB 1160 CCGCGGAGATGTTCTCTCCAGTTCAATTTTACCCTGTGTCTTGAATCCGAGCAATTCCT 1219
OY 1141 AAAAAGCAATTTTGGAGCCCTTGTGACATTAATACGATGACAGTCTGAAGAGCCAG 1200
DB 1220 AAAAAGCAATTTTGGAGCCCTTGTGACATTAATACGATGACAGTCTGAAGAGCCAG 1279
OY 1201 CTACAGAGCTGTGACATTAAGAGGAGCCGATTAAGCCGCTTGTAGAGATGTC 1260
DB 1280 CTACAGAGCTGTGACATTAAGAGGAGCCGATTAAGCCGCTTGTAGAGATGTC 1339
OY 1261 TGTGCTGTGTGTGATCTCTCTCGCTTCCCTTCCCTTTCGACAGCCACCACTAGTCTC 1320
DB 1340 TGTGCTGTGTGTGATCTCTCTCGCTTCCCTTCCCTTTCGACAGCCACCACTAGTCTC 1399
OY 1321 CTGCTGAAACATCTTCTTAACTTCAACCCAGACCATATTCGTGTGCAAGCTTCAAGTTTA 1380
DB 1400 CTGCTGAAACATCTTCTTAACTTCAACCCAGACCATATTCGTGTGCAAGCTTCAAGTTTA 1459
OY 1381 TTTCAACCCAGAAAGCTCAATTTTGTCTTCAACATTTGTGAAATTTCTGTACTGACACA 1440
DB 1460 TTTCAACCCAGAAAGCTCAATTTTGTCTTCAACATTTGTGAAATTTCTGTACTGACACA 1519
OY 1441 ACAGAGGTTCTGCGAAGGAGATGTAACAGGCTGGCTGGCTTGTGTGCTTCAAGTT 1500
DB 1520 ACAGAGGTTCTGCGAAGGAGATGTAACAGGCTGGCTGGCTTGTGTGCTTCAAGTT 1579
OY 1501 CTTCAGCCAAACATACATGATCCCATGAAAGACAGCGGAAAGCCCTGCTCTTAAGATA 1560
DB 1580 CTTCAGCCAAACATACATGATCCCATGAAAGACAGCGGAAAGCCCTGCTCTTAAGATA 1639
OY 1561 TCGATCTTCTCTGAAACAAATTTTCTCACTTACAGATGACCCCTCAATCCCATC 1620
DB 1640 TCGATCTTCTCTGAAACAAATTTTCTCACTTACAGATGACCCCTCAATCCCATC 1699
OY 1621 ATTAATGTTGGTCCAGGAACCCGCAATGACCCGTTAATTTGGGTTCTTACAACTAGAGAG 1680
DB 1700 ATTAATGTTGGTCCAGGAACCCGCAATGACCCGTTAATTTGGGTTCTTACAACTAGAGAG 1759
OY 1681 AAACCTCAAGAAACAACCCAGATGAAATTTTGAAGCAATGTGTTTGTGCTGC 1740
DB 1760 AAACCTCAAGAAACAACCCAGATGAAATTTTGAAGCAATGTGTTTGTGCTGC 1819
OY 1741 AGGCATTAAGATGAGATTAATCTATTCAAGAAAGAGCTCAAGATTTCTTAAGCATGGG 1800
DB 1820 AGGCATTAAGATGAGATTAATCTATTCAAGAAAGAGCTCAAGATTTCTTAAGCATGGG 1879
OY 1801 ATCTTAATCACTTAAGGTTTCTCTCAAGAGATGCTCTGTGAGGAGAGAGAGCC 1860
DB 1880 ATCTTAATCACTTAAGGTTTCTCTCAAGAGATGCTCTGTGAGGAGAGAGAGCC 1939
OY 1861 CCAGCAAGATATGTAACAAGCAACATCCAGTTCATGAGCCAGAGGTGGAGAAATCTCTC 1920
DB 1940 CCAGCAAGATATGTAACAAGCAACATCCAGTTCATGAGCCAGAGGTGGAGAAATCTCTC 1999
OY 1921 CTCACAGAGAGGCGCATATTTATGTGTGTGAGATGCAAGAAATATGCGCAAGATGTA 1980
DB 2000 CTCACAGAGAGGCGCATATTTATGTGTGTGAGATGCAAGAAATATGCGCAAGATGTA 2059
OY 1981 CATGATGCCCTTGTGAATAATTAAGCAAGAGGTTGAGTTGAAGAACTTGAAGCAATG 2040
DB 2060 CATGATGCCCTTGTGAATAATTAAGCAAGAGGTTGAGTTGAAGAACTTGAAGCAATG 2119
OY 2041 AAAACCTGCGCACTTTAAAGAGAAAGAAACGCTACCTTCAAGATATTTGTGATATA 2097
DB 2120 AAAACCTGCGCACTTTAAAGAGAAAGAAACGCTACCTTCAAGATATTTGTGATATA 2176
```

RESULT 8  
AD087538  
ID AD087538 standard; cDNA; 3270 BP.



XX ADQ87538;  
XX 07-OCT-2004 (first entry)  
XX  
DE Human tumour-associated antigenic target (TAT) cDNA sequence #4416.  
XX  
XX human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;  
XX cancer; cell proliferative disorder; gene; ss.  
XX Homo sapiens.  
XX MO2004060270-A2.  
XX  
XX 22-JUL-2004.  
XX  
XX 15-OCT-2003; 2003MO-US029126.  
XX  
XX 18-OCT-2002; 2002US-0418988P.  
XX  
XX (GETH ) GENENTECH INC.  
XX (WUTD/) WU T D.  
XX (ZHOU/) ZHOU Y.  
XX  
XX Mu TD, Zhou Y;  
XX  
XX WPI; 2004-534300/51.  
XX  
XX New nucleic acid molecule and encoded polypeptide, for diagnosing,  
XX preventing or treating cell proliferative disorders such as cancer.  
XX  
XX Claim 1; SEQ ID NO 4416; 5504bp; English.  
XX  
XX The present invention describes an isolated tumour-associated antigenic  
XX target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide  
XX sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of  
XX (a); (c) the complement of (a) or (b); (d) a sequence that has 80%  
XX sequence identity to (a)-(c); or (e) a sequence that hybridises to (a) -  
XX (c). Also described: (1) an expression vector comprising the above  
XX nucleic acid; (2) a host cell comprising the above expression vector; (3)  
XX a process for producing a polypeptide; (4) an isolated polypeptide  
XX comprising: (a) an amino acid sequence encoded by any of the above  
XX nucleotide sequences; (b) an amino acid sequence encoded by the full-  
XX length coding region of the above nucleotide sequences; or (c) a sequence  
XX having at least 80% identical to (a) or (b); (5) a chimeric polypeptide  
XX comprising the above polypeptide fused to a heterologous polypeptide; (6)  
XX an isolated antibody that binds to the above polypeptide; (7) a process  
XX for producing the antibody; (8) an isolated oligopeptide that binds to  
XX the above polypeptide; (9) a tumour-associated antigenic target (TAT)  
XX binding organic molecule that binds to the above polypeptide; (10) a  
XX composition of matter comprising the above (chimeric) polypeptide,  
XX antibody, oligopeptide or TAT binding organic molecule, in combination  
XX with a carrier; (11) an article of manufacture comprising a container and  
XX the composition of matter contained within the container; (12) methods of  
XX inhibiting the growth of a cell that expresses the above protein, where  
XX the growth of the cell is at least in part dependent upon a growth  
XX potential effect of the above protein; (13) a method of  
XX therapeutically treating a mammal having a cancerous tumour comprising  
XX cells that express the above protein; (14) a method of determining the  
XX presence of a protein in a sample suspected of containing the protein  
XX described above; (15) methods of diagnosing the presence of a tumour in a  
XX mammal; (16) a method for treating or preventing a cell proliferative  
XX disorder associated with increased expression or activity of the above  
XX protein; and (17) a method of binding an antibody, oligopeptide or  
XX organic molecule to a cell that expresses the protein described above.  
XX The TAT sequences have cytostatic activities and can be used in gene  
XX therapy. The composition and methods are useful for diagnosing,  
XX preventing or treating cancer. The composition is also used for preparing  
XX a medicament for the therapeutic treatment or diagnostic detection of a  
XX cell proliferative disorder or cancer. The present sequence represents a  
XX human TAT cDNA sequence from the present invention.  
XX  
XX Sequence 3270 BP; 934 A; 702 C; 680 G; 954 T; 0 U; 0 Other;

Query Match 90.3%; Score 1893; DB 13; Length 3270;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2093; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 1 ATGAGAGAGTTTCTGTACTATATGCTACACAGCAGGACAGGCAAAAGCCATCGCAGAA 60  
DB 112 ATGAGAGAGTTTCTGTACTATATGCTACACAGCAGGACAGGCAAAAGCCATCGCAGAA 171  
QY 61 GAAATGTGAGCAAGCTGTGTACATGATGATTTTCTGAGATCTTCACTATATTAGTAA 120  
DB 172 GAAATGTGAGCAAGCTGTGTACATGATGATTTTCTGAGATCTTCACTATATTAGTAA 231  
QY 121 TCCGATPAGTAGTACTTAAACCGAAGAGCTCTCTTGTGTGTGTTTCTACACG 180  
DB 232 TCCGATPAGTAGTACTTAAACCGAAGAGCTCTCTTGTGTGTGTTTCTACACG 291  
QY 181 GGCACCGGAGACCCAGCCGACAGCCGCAAGTTGTTAAGAAATPACAGAACCAACA 240  
DB 292 GGCACCGGAGACCCAGCCGACAGCCGCAAGTTGTTAAGAAATPACAGAACCAACA 351  
QY 241 CTGCGGTTGATTTCTTGTCTCAGCTGCGGTATGGTTACTGGGTCTCGGTGATTCAGAA 300  
DB 352 CTGCGGTTGATTTCTTGTCTCAGCTGCGGTATGGTTACTGGGTCTCGGTGATTCAGAA 411  
QY 301 TACACCTACTTTTGCAATGGGGGGAAGATPATTGATPAAAGACTTCAGAGCTTGAAGCC 360  
DB 412 TACACCTACTTTTGCAATGGGGGGAAGATPATTGATPAAAGACTTCAGAGCTTGAAGCC 471  
QY 361 CGGCAATTTTCATGACATGACATGACATGACATGACATGACATGACATGACATGACATG 420  
DB 472 CGGCAATTTTCATGACATGACATGACATGACATGACATGACATGACATGACATGACATG 531  
QY 421 CCGTGATTTGCTGAGCTTGGCCAGCCCTCAGAAACATTTTATGATCAAGCAGAGCA 480  
DB 532 CCGTGATTTGCTGAGCTTGGCCAGCCCTCAGAAACATTTTATGATCAAGCAGAGCA 591  
QY 481 GAGAGATPAGTGGCGCACTCCGGTGGCATCCTGATCCTTGAAGAGACCTTGTG 540  
DB 592 GAGAGATPAGTGGCGCACTCCGGTGGCATCCTGATCCTTGAAGAGACCTTGTG 651  
QY 541 AAGTCAGAGCTGTACATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 600  
DB 652 AAGTCAGAGCTGTACATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 711  
QY 601 AGAAGAGATTTGAGGTTTGAAGCAAAATGACAGTAAAGCAACCAATCCATGTTGTA 660  
DB 712 AGAAGAGATTTGAGGTTTGAAGCAAAATGACAGTAAAGCAACCAATCCATGTTGTA 771  
QY 661 ATTGAAGACTTGAATCTCTCACTTACCCGTTGGTACCCCACTTCACAGCTCTCTG 720  
DB 772 ATTGAAGACTTGAATCTCTCACTTACCCGTTGGTACCCCACTTCACAGCTCTCTG 831  
QY 721 AATATTCCTGTTTACCCCGAATTTTACAGGTACATCTGAGAGAGTCTCTGGCCAG 780  
DB 832 AATATTCCTGTTTACCCCGAATTTTACAGGTACATCTGAGAGAGTCTCTGGCCAG 891  
QY 781 GAGGAAGCAAGTATCTGTGACTTCAAGAGATCCAGTTTTCATGAGCAATTTCAAG 840  
DB 892 GAGGAAGCAAGTATCTGTGACTTCAAGAGATCCAGTTTTCATGAGCAATTTCAAG 951  
QY 841 GCAATTCACCTTACATGAAATGATGCAATPAAACCACTCTGCTGTGATGATTTGACAT 900  
DB 952 GCAATTCACCTTACATGAAATGATGCAATPAAACCACTCTGCTGTGATGATTTGACAT 1011  
QY 901 TCMAATACAGACTTTTCTCATGAGCTTGAAGATGCTTCAAGCTATCTGCTTAAACAGT 960  
DB 1012 TCMAATACAGACTTTTCTCATGAGCTTGAAGATGCTTCAAGCTATCTGCTTAAACAGT 1071  
QY 961 GATTTCGAGTACAAAGCTTCCAAAGAGTGCAGCTTGAAGATPAAAGAGAGAGAGAGAGAG 1020  
DB 1072 GATTTCGAGTACAAAGCTTCCAAAGAGTGCAGCTTGAAGATPAAAGAGAGAGAGAGAG 1131

QY 1021 GTCTTTGAAATAAAGCAGACACAAGAGAGAGCTTACCTTACCAGCATATA 1080  
 DB 1132 GTCTTTGAAATAAAGCAGACACAAGAGAGAGCTTACCTTACCAGCATATA 1191  
 QY 1081 CTTGGGGGATGTTCTCTCAGATTCAATTTTACCTGGTGTCTTGAATCCGAGCAATTCCT 1140  
 DB 1192 CTTGGGGGATGTTCTCTCAGATTCAATTTTACCTGGTGTCTTGAATCCGAGCAATTCCT 1251  
 QY 1141 AAAAAGGCAATTTTGGAGCCCTTGGAGCTATACAGAGAGAGCTGAGAAAAGGCGAGG 1200  
 DB 1252 AAAAAGGCAATTTTGGAGCCCTTGGAGCTATACAGAGAGAGCTGAGAAAAGGCGAGG 1311  
 QY 1201 CTACAGAGAGCTGAGTAAACAAGAGGAGCGAGTATATAGCCGCTTTGTATGAGATGCC 1260  
 DB 1312 CTACAGAGAGCTGAGTAAACAAGAGGAGCGAGTATATAGCCGCTTTGTATGAGATGCC 1371  
 QY 1261 TGTGCTGCTTGTGGATCTCTCTCGCTTTCCTTCTTGGCAGCAGCACTCAAGTCTC 1320  
 DB 1372 TGTGCTGCTTGTGGATCTCTCTCGCTTTCCTTCTTGGCAGCAGCACTCAAGTCTC 1431  
 QY 1321 CTGCTGGAACATCTTCTTAACTTCAACCCAGACATATTCGATGCAAGCTCAAGTTTA 1380  
 DB 1432 CTGCTGGAACATCTTCTTAACTTCAACCCAGACATATTCGATGCAAGCTCAAGTTTA 1491  
 QY 1381 TTTCAACCCAGAGAGCTCCATTTTGTCTTCAACATTTGGAATTTCTGTCTACTGCGACA 1440  
 DB 1492 TTTCAACCCAGAGAGAGCTCCATTTTGTCTTCAACATTTGGAATTTCTGTCTACTGCGACA 1551  
 QY 1441 ACAGAGGTTCTGCGAGAGAGATGATACAGGCTGCTGCTTGTGGTTCCTTCAAGTT 1500  
 DB 1552 ACAGAGGTTCTGCGAGAGAGATGATGATACAGGCTGCTGCTTGTGGTTCCTTCAAGTT 1611  
 QY 1501 CTTGAGCCGAACATCATGATCCCATGAGAGACAGCGGAGAAAGCCCTGCTCTTAAGATA 1560  
 DB 1612 CTTGAGCCGAACATCATGATCCCATGAGAGACAGCGGAGAAAGCCCTGCTCTTAAGATA 1671  
 QY 1561 TCCATCTCTCTCTGAGACACAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
 DB 1672 TCCATCTCTCTCTGAGACACAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1731  
 QY 1621 ATATGTTGGTTCAGAGAACCGGAGATAGCCCGTTTATTTGGTTCCTTCAACATAGAGAG 1680  
 DB 1732 ATATGTTGGTTCAGAGAACCGGAGATAGCCCGTTTATTTGGTTCCTTCAACATAGAGAG 1791  
 QY 1681 AAACCTCCAGAGACACACCCAGATGGAATTTTGGAGCAATGTGTTGTTTGGCTGTC 1740  
 DB 1792 AAACCTCCAGAGACACACCCAGATGGAATTTTGGAGCAATGTGTTGTTTGGCTGTC 1851  
 QY 1741 AGGCAATAGGATAGGATATATCTATTCAGAAAAAGAGCTCAGACATTTCTTAAAGATGGG 1800  
 DB 1852 AGGCAATAGGATAGGATATATCTATTCAGAAAAAGAGCTCAGACATTTCTTAAAGATGGG 1911  
 QY 1801 ATCTTAATCTCTTAAAGGTTTCTTCTGAGAGATGCTCTGTTGGGAGAGAGAGAGCC 1860  
 DB 1912 ATCTTAATCTCTTAAAGGTTTCTTCTGAGAGATGCTCTGTTGGGAGAGAGAGAGCC 1971  
 QY 1861 CAGAGAAATATATGACAGACACATCCAGCTTCATGCGCAGAGAGGTGGCGAATCTCTC 1920  
 DB 1972 CAGAGAAATATATGACAGACACATCCAGCTTCATGCGCAGAGAGGTGGCGAATCTCTC 2031  
 QY 1921 CTCAGAGAGAGGCGCATTTATGTTGTTGAGATGAGCAAGAAATATGCGCAAGATGTA 1980  
 DB 2032 CTCAGAGAGAGGCGCATTTATGTTGTTGAGATGAGCAAGAAATATGCGCAAGATGTA 2091  
 QY 1981 CATGATGCCCTTGTGCAATTAATAGCAAGAGGTTGAGTTGAAAAAATAGAGCAATG 2040  
 DB 2092 CATGATGCCCTTGTGCAATTAATAGCAAGAGGTTGAGTTGAAAAAATAGAGCAATG 2151  
 QY 2041 AAAACCTTGGCCACTTTAAAAAGAGAAAGAGTACTTCAAGATATTTGGTCTATA 2097  
 DB 2152 AAAACCTTGGCCACTTTAAAAAGAGAAAGAGTACTTCAAGATATTTGGTCTATA 2208

RESULT 9  
 ADM43216  
 ID ADM43216 standard; cDNA; 2091 BP.  
 XX  
 AC ADM43216;  
 XX  
 DT 03-JUN-2004 (first entry)  
 XX  
 DE Human methionine synthase reductase CDS del 1726-1728 variant.  
 XX  
 KW Human; ss; Methionine synthase reductase polypeptide; HsMTRR; cancer;  
 KW cardiovascular disease; neural tube defect; hyperhomocysteinemia;  
 XX chromosome 5p15.2-p15.3; SNP; single nucleotide polymorphism.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT 1..2091  
 FT CDS /tag= a  
 FT /product= "HsMTRRdelR559"  
 FT /partial  
 FT /note= "No stop codon shown"  
 FT variation replace(66,A)  
 FT /tag= b  
 FT /standard\_name= "Single\_nucleotide\_polymorphism"  
 FT variation replace(110,A)  
 FT /tag= c  
 FT /standard\_name= "Single\_nucleotide\_polymorphism"  
 FT variation replace(1726,TTGT)  
 FT /tag= d  
 PN US2003082676-A1.  
 PD 01-MAY-2003.  
 XX  
 PF 10-AUG-1999; 99US-00371347.  
 XX  
 PR 16-JAN-1998; 98US-0071622P.  
 PR 15-JAN-1999; 99US-00232028.  
 XX  
 PA (GRAV/) GRAVEL R A.  
 PA (ROZE/) ROZEN R.  
 PA (LECL/) LECLERC D.  
 PA (WILS/) WILSON A.  
 PA (ROSE/) ROSENBLATT D.  
 PI Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;  
 DR MPI; 2003-576610/54.  
 DR P-PSDB; ADM43217.  
 XX  
 PT New substantially pure nucleic acid encoding a mammalian methionine  
 PT synthase reductase polypeptide, useful for diagnosing, preventing or  
 PT treating conditions associated with altered methionine synthase activity,  
 PT e.g. cancer.  
 XX  
 PS Disclosure; SEQ ID NO 45; 26pp; English.  
 XX  
 CC The invention relates to a substantially pure nucleic acid that encodes a  
 CC mammalian methionine synthase reductase polypeptide, HsMTRR, or that  
 CC hybridizes at high stringency to a nucleic acid appearing as ADM43208 or  
 CC ADM43209. Also included are a non-human animal where one or both genetic  
 CC alleles encoding the methionine synthase reductase polypeptide are  
 CC mutated, an antibody that specifically binds the above methionine  
 CC synthase reductase polypeptide, a method of detecting the presence of the  
 CC above polypeptide, a method for detecting sequence variants for  
 CC methionine synthase reductase in a mammal, methods of treating or  
 CC preventing cancer (or cardiovascular disease or neural tube defects) in a  
 CC subject, methods of screening for a compound that modulates methionine  
 CC synthase reductase biological activity and a method for detecting an  
 CC increased risk of developing a neural tube defect in a mammalian embryo  
 CC or foetus. The nucleic acid is useful in diagnosing, preventing or  
 CC treating conditions associated with altered methionine synthase activity,

CC such as cancer, cardiovascular disease or neural tube defects, or in  
CC screening for a compound that modulates methionine synthase reductase  
CC biological activity. Naturally occurring variants of the polypeptide are  
CC also associated with hyperhomocysteinemia. The gene for HmTR is  
CC located on chromosome 5p15.2-p15.3. The present sequence is the coding  
CC sequence of a variant human hsmTR cDNA.

XX Sequence 2091 BP; 591 A; 489 C; 480 G; 531 T; 0 U; 0 Other;

Query Match 85.8%; Score 1800; DB 11; Length 2091;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2090; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 1 ATGAGAGGTTTCTGTTACTATATGCTACACAGCAGGACAGGCAAGGCCATCGAGAA 60  
DB 1 ATGAGAGGTTTCTGTTACTATATGCTACACAGCAGGACAGGCAAGGCCATCGAGAA 60  
QY 61 GAAATGTGAGAGAGCTGTGTATCATGTATTTCTGAGATCTTCATCTATATTAGGAA 120  
DB 61 GAAATGTGAGAGAGCTGTGTATCATGTATTTCTGAGATCTTCATCTATATTAGGAA 120  
QY 121 TCCGATTAATGATACCTTAACCAAGCAGCTCTCTGTGTGTGTGTCTTACACG 180  
DB 121 TCCGATTAATGATACCTTAACCAAGCAGCTCTCTGTGTGTGTGTCTTACACG 180  
QY 181 GGCACCGAGACCCACCCGACACAGCCCGAAGTTGTAAAGAAATACAAACCA 240  
DB 181 GGCACCGAGACCCACCCGACACAGCCCGAAGTTGTAAAGAAATACAAACCA 240  
QY 241 CTGCGGTTGATTTCTTTGTCTACCTGTGGGTATGGGTACTGGGTCTCGGTGATTCAGAA 300  
DB 241 CTGCGGTTGATTTCTTTGTCTACCTGTGGGTATGGGTACTGGGTCTCGGTGATTCAGAA 300  
QY 301 TACACCTACTTTTGCAATGGGGGAGATATTGATTAAGCACTTCAAGACTTGGAGCC 360  
DB 301 TACACCTACTTTTGCAATGGGGGAGATATTGATTAAGCACTTCAAGACTTGGAGCC 360  
QY 361 CGGCATTTCTATGACACTGACATGACATGACTGTGTAGGTTTGAACCTTGTGTGAG 420  
DB 361 CGGCATTTCTATGACACTGACATGACATGACTGTGTAGGTTTGAACCTTGTGTGAG 420  
QY 421 CCGTGAATGCTGTGACTGTGGCCAGCCCTCAGAAAGATTTTATAGTCAACAGAGGACA 480  
DB 421 CCGTGAATGCTGTGACTGTGGCCAGCCCTCAGAAAGATTTTATAGTCAACAGAGGACA 480  
QY 481 GAGAGATTAAGTGGCGCACTCCCGGTGGCATCTGCACTTGTAGGAGACAGACCTTGTG 540  
DB 481 GAGAGATTAAGTGGCGCACTCCCGGTGGCATCTGCACTTGTAGGAGACAGACCTTGTG 540  
QY 541 AAGTCAGAGCTGTACACATTTGAATTCGAAGTCGAGCTTGTGAGATTGATTCAGGA 600  
DB 541 AAGTCAGAGCTGTACACATTTGAATTCGAAGTCGAGCTTGTGAGATTGATTCAGGA 600  
QY 601 AGAAGGATTTCTGAGTTTGAAGCAAAATGCAGTGAACAGCAACCAATCCAATGTGTGA 660  
DB 601 AGAAGGATTTCTGAGTTTGAAGCAAAATGCAGTGAACAGCAACCAATCCAATGTGTGA 660  
QY 661 ATTGAAGACTTGTGCTCTCACTTACCCGTTCCGTATCCCACTCTCAAGCCTCTCTG 720  
DB 661 ATTGAAGACTTGTGCTCTCACTTACCCGTTCCGTATCCCACTCTCAAGCCTCTCTG 720  
QY 721 AATATTCCTGTGTTACCCCGAATATTATACAGTATCATCTGCGAGAGTCTCTTGGCCAG 780  
DB 721 AATATTCCTGTGTTACCCCGAATATTATACAGTATCATCTGCGAGAGTCTCTTGGCCAG 780  
QY 781 GAGGAAGCCAAAGATCTGTGACTTCAGAGATTCAGATTTTCAAGGCAATTTCAAG 840  
DB 781 GAGGAAGCCAAAGATCTGTGACTTCAGAGATTCAGATTTTCAAGGCAATTTCAAG 840  
QY 841 GCATTTCACTTACGAAATGATGCAATAAAACCACTCTGCTGTGATGATTTGCAATT 900  
DB 841 GCATTTCACTTACGAAATGATGCAATAAAACCACTCTGCTGTGATGATTTGCAATT 900

QY 901 TCMAATACAGACTTTTCTTATCAGCCTGAGATGCTTACGGTGTATCTGCCCTAACAT 960  
DB 901 TCMAATACAGACTTTTCTTATCAGCCTGAGATGCTTACGGTGTATCTGCCCTAACAT 960  
QY 961 GATTCTGAGGTACAAAGCTTACTCCAAAGACTGCACTTGAAGATTAAGAGACACTGC 1020  
DB 961 GATTCTGAGGTACAAAGCTTACTCCAAAGACTGCACTTGAAGATTAAGAGACACTGC 1020  
QY 1021 GTCTTTTGAATAAAGGACAGACAAAGAAAGAGCTACCTTACCCGACATATA 1080  
DB 1021 GTCTTTTGAATAAAGGACAGACAAAGAAAGAGCTACCTTACCCGACATATA 1080  
QY 1081 CCGCGGGATGTTCTCTCCAGTTCAATTTTACTGTGTCTTGAATTCGAGCAATTCCT 1140  
DB 1081 CCGCGGGATGTTCTCTCCAGTTCAATTTTACTGTGTGTCTTGAATTCGAGCAATTCCT 1140  
QY 1141 AAAAAGCAATTTTGGAGCCCTTGTGAGCTATACAGGACAGTCTGAAAAAGCCGAG 1200  
DB 1141 AAAAAGCAATTTTGGAGCCCTTGTGAGCTATACAGGACAGTCTGAAAAAGCCGAG 1200  
QY 1201 CTACAGAGCTGTGACGTAAACAAAGGGGACCGATTATAGCCGTTTGTACGAGATGCC 1260  
DB 1201 CTACAGAGCTGTGACGTAAACAAAGGGGACCGATTATAGCCGTTTGTACGAGATGCC 1260  
QY 1261 TGTGCTGCTGTGTGATCTCTCTGCTGCTTCCCTTCTTCCAGGACCACTCAATCTC 1320  
DB 1261 TGTGCTGCTGTGTGATCTCTCTGCTGCTTCCCTTCTTCCAGGACCACTCAATCTC 1320  
QY 1321 CTGCTGAAATCTTCTTAACTTCAACCCAGACATATTCGTGTGCAAGCTCAAGTTTA 1380  
DB 1321 CTGCTGAAATCTTCTTAACTTCAACCCAGACATATTCGTGTGCAAGCTCAAGTTTA 1380  
QY 1381 TTTCAACCAAGAAAGCTTCATTTTGTCTTCAACATTTGSGAATTTCTGTCTACTGCGACA 1440  
DB 1381 TTTCAACCAAGAAAGCTTCATTTTGTCTTCAACATTTGSGAATTTCTGTCTACTGCGACA 1440  
QY 1441 ACAGAGTTCTGTGGGAAAGGATATGATCAGGCTGGCTGCTGTGTGTGTTCAAGTT 1500  
DB 1441 ACAGAGTTCTGTGGGAAAGGATATGATCAGGCTGGCTGCTGTGTGTGTTCAAGTT 1500  
QY 1501 CTTGACGCAAACTACATGATCCATGAAAGACAGGGGAAAGCCCTGAGCTCTAAGATA 1560  
DB 1501 CTTGACGCAAACTACATGATCCATGAAAGACAGGGGAAAGCCCTGAGCTCTAAGATA 1560  
QY 1561 TCCATCTCTCTGGAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
DB 1561 TCCATCTCTCTGGAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
QY 1621 ATATATGTTGGTTCAGAAACCGGATAGCCCGTTTATTTGGGTTCTTACAACTAAGAG 1680  
DB 1621 ATATATGTTGGTTCAGAAACCGGATAGCCCGTTTATTTGGGTTCTTACAACTAAGAG 1680  
QY 1681 AAATCTCAAGAAACAAACCCAGATGGAATTTTGAAGCAATGTGTGTTTGTGCTGC 1740  
DB 1681 AAATCTCAAGAAACAAACCCAGATGGAATTTTGAAGCAATGTGTGTTTGTGCTGC 1740  
QY 1741 AGGCATTAAGATTAAGATTAATCTATTCAAGAAAGCTCAGACATTTCTTAAGATGG 1800  
DB 1741 AGGCATTAAGATTAAGATTAATCTATTCAAGAAAGCTCAGACATTTCTTAAGATGG 1800  
QY 1738 AGGCATTAAGATTAAGATTAATCTATTCAAGAAAGCTCAGACATTTCTTAAGATGG 1797  
DB 1738 AGGCATTAAGATTAAGATTAATCTATTCAAGAAAGCTCAGACATTTCTTAAGATGG 1797  
QY 1801 ATCTTAATCATTAAGGTTTCTTCTCAAGAGATGCTCTCTGTGGGAGAGGAAAGCC 1860  
DB 1801 ATCTTAATCATTAAGGTTTCTTCTCAAGAGATGCTCTCTGTGGGAGAGGAAAGCC 1860  
QY 1798 ATCTTAATCATTAAGGTTTCTTCTCAAGAGATGCTCTCTGTGGGAGAGGAAAGCC 1857  
DB 1798 ATCTTAATCATTAAGGTTTCTTCTCAAGAGATGCTCTCTGTGGGAGAGGAAAGCC 1857  
QY 1861 CCAGCAAGATTAATACAGCAACATCCAGCTTCAAGGACAGAGGTGGGCAATCTCTC 1920  
DB 1861 CCAGCAAGATTAATACAGCAACATCCAGCTTCAAGGACAGAGGTGGGCAATCTCTC 1920  
QY 1858 CCAGCAAGATTAATACAGCAACATCCAGCTTCAAGGACAGAGGTGGGCAATCTCTC 1917  
DB 1858 CCAGCAAGATTAATACAGCAACATCCAGCTTCAAGGACAGAGGTGGGCAATCTCTC 1917  
QY 1921 CTCACAGAGAACGGCCATATTTATGTGTGTGAGATGCAAGAAATATGCGCAAGATGTA 1980  
DB 1921 CTCACAGAGAACGGCCATATTTATGTGTGTGAGATGCAAGAAATATGCGCAAGATGTA 1980  
QY 1918 CTCACAGAGAACGGCCATATTTATGTGTGTGAGATGCAAGAAATATGCGCAAGATGTA 1977  
DB 1918 CTCACAGAGAACGGCCATATTTATGTGTGTGAGATGCAAGAAATATGCGCAAGATGTA 1977  
QY 1981 CATGATGCCCTTGTGCAAAATATTAAGCAAGAGGTTGAGTTGAAAACTAAGAACATG 2040

|||||  
DB 1978 CATGATGCCCTTGCAATTAATTAACCAAGAGTTGAGTTGAAAACTAGAACGAATG 2037  
QY 2041 AAAACCCCTGCGCACTTTAAAAAGAAAGAAACCGTACTCTCAGATATTTGGTCA 2094  
DB 2038 AAAACCTGCGCACTTTAAAAAGAAAGAAACCGTACTCTCAGATATTTGGTCA 2091  
RESULT 10  
ADM43214  
ID ADM43214 standard; cDNA; 2091 BP.  
XX  
AC ADM43214;  
XX  
DT 03-JUN-2004 (first entry)  
XX  
DE Human methionine synthase reductase CDS del 1675-1678 variant.  
XX  
KW Human; ss; Methionine synthase reductase polypeptide; HsMTRR; cancer;  
KM cardiovascular disease; neural tube defect; hyperhomocysteinemia;  
XX chromosome 5p15.2-p15.3; SNP; single nucleotide polymorphism.  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..2091  
FT /tag= a  
FT /product= "HsMTRRdelR559"  
FT /partial  
FT /note= "No stop codon shown"  
FT /replace(66,A)  
FT /tag= b  
FT /standard name= "single\_nucleotide polymorphism"  
FT /replace(110,A)  
FT /tag= c  
FT /standard name= "single\_nucleotide polymorphism"  
FT /replace(1675,AGAG)  
FT /tag= d  
XX  
PN US2003082676-A1.  
XX  
XX 01-MAY-2003.  
XX  
PF 10-AUG-1999; 99US-00371347.  
XX  
PR 16-JAN-1998; 98US-0071622P.  
PR 15-JAN-1999; 99US-00232028.  
XX  
PA (GRAV/) GRAVEL R. A.  
PA (ROZE/) ROZEN R.  
PA (LECL/) LECLEERC D.  
PA (WILS/) WILSON A.  
PA (ROSE/) ROSENBLATT D.  
XX  
PI Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;  
XX WPI; 2003-576610/54.  
XX P-PSDB; ADM43215.  
XX  
XX New substantially pure nucleic acid encoding a mammalian methionine  
PT synthase reductase polypeptide, useful for diagnosing, preventing or  
PT treating conditions associated with altered methionine synthase activity,  
PT e.g. cancer.  
XX  
PS Disclosure; SEQ ID NO 47; 26bp; English.  
XX  
XX The invention relates to a substantially pure nucleic acid that encodes a  
CC mammalian methionine synthase reductase polypeptide, HsMTRR, or that  
CC hybridizes at high stringency to a nucleic acid appearing as ADM43208 or  
CC ADM43209. Also included are a non-human animal where one or both genetic  
CC alleles encoding the methionine synthase reductase polypeptide are  
CC mutated, an antibody that specifically binds the above methionine  
CC synthase reductase polypeptide, a method of detecting the presence of the

CC above polypeptide, a method for detecting sequence variants for  
CC methionine synthase reductase in a mammal, methods of treating or  
CC preventing cancer (or cardiovascular disease or neural tube defects) in a  
CC subject, methods of screening for a compound that modulates methionine  
CC synthase reductase biological activity and a method for detecting an  
CC increased risk of developing a neural tube defect in a mammalian embryo  
CC or fetus. The nucleic acid is useful in diagnosing, preventing or  
CC treating conditions associated with altered methionine synthase activity,  
CC such as cancer, cardiovascular disease or neural tube defects, or in  
CC screening for a compound that modulates methionine synthase reductase  
CC biological activity. Naturally occurring variants of the polypeptide are  
CC also associated with hyperhomocysteinemia. The gene for HsMTRR is  
CC located on chromosome 5p15.2-p15.3. The present sequence is the coding  
CC sequence of a variant human hsmtrr cDNA.  
XX

Seq Sequence 2091 BP; 589 A; 489 C; 480 G; 533 T; 0 U; 0 Other;

Query Match 85.8%; Score 1800; DB 11; Length 2091;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 2090; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 1 ATGAGGAGGTTCTGTACTATATGCTACACAGGAGGACAGGCAAGGCCATCGAGAA 60  
DB 1 ATGAGGAGGTTCTGTACTATATGCTACACAGGAGGACAGGCAAGGCCATCGAGAA 60  
QY 61 GAAATGTGAGACAGCTGTGTAATGATGTTTCTGAGATCTTCACTATATTAGTAA 120  
DB 61 GAAATGTGAGACAGCTGTGTAATGATGTTTCTGAGATCTTCACTATATTAGTAA 120  
QY 121 TCGATTAAGTATGACTTAAACCGAAACAGCTCTTGTGTTGTGTTCTACACG 180  
DB 121 TCGATTAAGTATGACTTAAACCGAAACAGCTCTTGTGTTGTGTTCTACACG 180  
QY 121 TCGATTAAGTATGACTTAAACCGAAACAGCTCTTGTGTTGTGTTCTACACG 180  
DB 121 TCGATTAAGTATGACTTAAACCGAAACAGCTCTTGTGTTGTGTTCTACACG 180  
QY 181 GGCACCGGAGACCCACCCACACAGCCGCAAGTTGTTAAAGAAATACGAACCAACA 240  
DB 181 GGCACCGGAGACCCACCCACACAGCCGCAAGTTGTTAAAGAAATACGAACCAACA 240  
QY 241 CTGCCGTTGATTTCTTTCCTCACCCTGCGGTATGAGTTCTGGGTCTCGGTGATTCAGAA 300  
DB 241 CTGCCGTTGATTTCTTTCCTCACCCTGCGGTATGAGTTCTGGGTCTCGGTGATTCAGAA 300  
QY 301 TACACCTACTTTTGCATATGAGGAGGAAATATATGATTAACGACTTCAGAGCTTGAGCC 360  
DB 301 TACACCTACTTTTGCATATGAGGAGGAAATATATGATTAACGACTTCAGAGCTTGAGCC 360  
QY 361 CGGCAATTTATGACACTGACATGCAATGACTGTGTGTTAGAACTTGTGTTGAG 420  
DB 361 CGGCAATTTATGACACTGACATGCAATGACTGTGTGTTAGAACTTGTGTTGAG 420  
QY 421 CCGTGATTTGCTGACCTCGGCGACGCCCTCAAGAACATTTTAGAGTCAAGAGAGACAA 480  
DB 421 CCGTGATTTGCTGACCTCGGCGACGCCCTCAAGAACATTTTAGAGTCAAGAGAGACAA 480  
QY 481 GAGAGATTAAGTGGGCACTCCCGTGATCATCTGACTCTTGAAGGACAGACTTGTG 540  
DB 481 GAGAGATTAAGTGGGCACTCCCGTGATCATCTGACTCTTGAAGGACAGACTTGTG 540  
QY 541 AAGTCAGAGCTGCTACACTTGAATTCAGATGAGCTTCTGAGATTCAGATTCAGGA 600  
DB 541 AAGTCAGAGCTGCTACACTTGAATTCAGATGAGCTTCTGAGATTCAGATTCAGGA 600  
QY 601 AGAAGGATTCGAGGTTTGAAGCAAAATGACAGTGAACAGCAACCAATCCATGTTGTA 660  
DB 601 AGAAGGATTCGAGGTTTGAAGCAAAATGACAGTGAACAGCAACCAATCCATGTTGTA 660  
QY 661 ATTGAAGACTTTGAGTCTCACTTACCCGTTGATACCCCACTTCACAAGCTCTGTG 720  
DB 661 ATTGAAGACTTTGAGTCTCACTTACCCGTTGATACCCCACTTCACAAGCTCTGTG 720  
QY 721 AATATTCCTGTTTACCCCAAGATTTTACAGTATCTGCAAGAGTCTTGTGCGCAG 780  
DB 721 AATATTCCTGTTTACCCCAAGATTTTACAGTATCTGCAAGAGTCTTGTGCGCAG 780

QY 781 GAGGAAAGCCAGATCTGTGACTTCAGCAGATCCAGTTTTCAGAGTCCAAATTTCAAG 840  
 DB 781 GAGGAAAGCCAGATCTGTGACTTCAGCAGATCCAGTTTTCAGAGTCCAAATTTCAAG 840  
 QY 841 GCGATTCATTTTCTAGAGTATGATGCGATTAAGCACTCTGCTGGTGAATTTGGACAT 900  
 DB 841 GCGATTCATTTTCTAGAGTATGATGCGATTAAGCACTCTGCTGGTGAATTTGGACAT 900  
 QY 901 TCAATTCAGACTTTTCTATTCAGCCTGAGATGCTTCAGCGTGAATCTGCTTAACT 960  
 DB 901 TCAATTCAGACTTTTCTATTCAGCCTGAGATGCTTCAGCGTGAATCTGCTTAACT 960  
 QY 961 GATTCTGAGGTACAAAGCCTTACTCAGAGCTGAGCTGAGATTAAGAGAGCACTGC 1020  
 DB 961 GATTCTGAGGTACAAAGCCTTACTCAGAGCTGAGCTGAGATTAAGAGAGCACTGC 1020  
 QY 1021 GTGCTTTTGAATTAAGGAGACAGCAAGAAAGAGAGCTTACCCGAGCAATTA 1080  
 DB 1021 GTGCTTTTGAATTAAGGAGACAGCAAGAAAGAGAGCTTACCCGAGCAATTA 1080  
 QY 1081 CCGTGGGATGTTCTCTCAATTCATTTTACCTGCTGTGAAATCCGAGCAATTCCT 1140  
 DB 1081 CCGTGGGATGTTCTCTCAATTCATTTTACCTGCTGTGAAATCCGAGCAATTCCT 1140  
 QY 1141 AAAAGGCAATTTTTCGAGGCTTGTGACTATACAGTGAAGAGTGTGAAAAGCCGAG 1200  
 DB 1141 AAAAGGCAATTTTTCGAGGCTTGTGACTATACAGTGAAGAGTGTGAAAAGCCGAG 1200  
 QY 1201 CTACAGAGCTGTGCACTTAACAAAGGGGAGCCGATTAATGAGCGCTTTGTAACAGATGCC 1260  
 DB 1201 CTACAGAGCTGTGCACTTAACAAAGGGGAGCCGATTAATGAGCGCTTTGTAACAGATGCC 1260  
 QY 1261 TGTGCTGCTGTGTGATCTCTCTCTGCTTCCCTTCTTGCGAGCAGCACTCACTC 1320  
 DB 1261 TGTGCTGCTGTGTGATCTCTCTCTGCTTCCCTTCTTGCGAGCAGCACTCACTC 1320  
 QY 1321 CTGCTGCAATCTTCTTAACTTCAACCAAGACATATTCGTGTGCAAGCTCAAGTTA 1380  
 DB 1321 CTGCTGCAATCTTCTTAACTTCAACCAAGACATATTCGTGTGCAAGCTCAAGTTA 1380  
 QY 1381 TTTTCCAGGAGAAAGCTTCAATTTGTCTGAACATTTGTAATTTCTGTCTACCTGACCA 1440  
 DB 1381 TTTTCCAGGAGAAAGCTTCAATTTGTCTGAACATTTGTAATTTCTGTCTACCTGACCA 1440  
 QY 1441 ACAAGGTTCTGCGAGAGGATGTATGATGAGCTGCTGCTGTGTGCTTCACT 1500  
 DB 1441 ACAAGGTTCTGCGAGAGGATGTATGATGAGCTGCTGCTGTGTGCTTCACT 1500  
 QY 1501 CTTGAGCCAAACATACATGATGCTCCATGAAAGACGCGGAAAGCCCTGCTCCCTAAGATA 1560  
 DB 1501 CTTGAGCCAAACATACATGATGCTCCATGAAAGACGCGGAAAGCCCTGCTCCCTAAGATA 1560  
 QY 1561 TCCATCTCTCTGCAAGCAAAATTTCTTCACTTACAGATGAGCCCTCAATCCCATC 1620  
 DB 1561 TCCATCTCTCTGCAAGCAAAATTTCTTCACTTACAGATGAGCCCTCAATCCCATC 1620  
 QY 1621 ATAATGTTGGTCCAGAGACCGGATGAGCCCGTTTATTTGGGTTCTTACAAATAGAG 1680  
 DB 1621 ATAATGTTGGTCCAGAGACCGGATGAGCCCGTTTATTTGGGTTCTTACAAATAGAG 1680  
 QY 1681 AAATCCAGAAACCAACCCAGATGAAATTTTGAGCAATGTGTTTGTGCTGC 1740  
 DB 1681 AAATCCAGAAACCAACCCAGATGAAATTTTGAGCAATGTGTTTGTGCTGC 1740  
 QY 1741 AGGATTAAGATAGGATTAATCTATTTGAGAAAGAGCTGAGCACTTTTCTTAAGCAGTGG 1800  
 DB 1741 AGGATTAAGATAGGATTAATCTATTTGAGAAAGAGCTGAGCACTTTTCTTAAGCAGTGG 1800  
 QY 1801 ATCTTAATCACTTAAGAGTTTCTTCTCAAGAGATGCTCTGTGTTGGGAGAGAGAGCC 1860  
 DB 1801 ATCTTAATCACTTAAGAGTTTCTTCTCAAGAGATGCTCTGTGTTGGGAGAGAGAGCC 1860  
 QY 1861 CCAGCAAAAGTATGTACAAAGACATCCAGCTTCAAGGAGAGGAGAGATCTC 1920

DB 1858 CCAGCAAAATATGTACAAAGCAATCCAGCTTCAATGCGACAGAGTGGCGAATCTC 1917  
 QY 1921 CTCGAGAGAGAGGCGCATTTTATGTGTGAGATGCAAGAAATATGAGCAAGATGTA 1980  
 DB 1918 CTCGAGAGAGAGGCGCATTTTATGTGTGAGATGCAAGAAATATGAGCAAGATGTA 1977  
 QY 1981 CATGATGCTCTGTGCAATTAATAGCAAGAGGTTGAGATTGAAAACTAAGCAATG 2040  
 DB 1978 CATGATGCTCTGTGCAATTAATAGCAAGAGGTTGAGATTGAAAACTAAGCAATG 2037  
 QY 2041 AAAACCTGCGCACTTTAAAGAGAAAGAGCTACCTCAGATATTGTCTCA 2094  
 DB 2038 AAAACCTGCGCACTTTAAAGAGAAAGAGCTACCTCAGATATTGTCTCA 2091

RESULT 11  
 AA58977  
 ID AA58977 standard; DNA; 3256 BP.  
 AC AA58977;  
 XX  
 DT 07-NOV-2000 (first entry)  
 XX  
 DE A human methionine synthase reductase DNA sequence with polymorphism.  
 XX  
 KW Human; methionine synthase reductase; MTRR; cancer;  
 KW cardiovascular disease; Down's Syndrome; neural tube defect;  
 KW premature coronary artery disease; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200042196-A2.  
 XX  
 PD 20-JUL-2000.  
 XX  
 PF 14-JAN-2000; 2000NC-IB000209.  
 XX  
 PR 15-JAN-1999; 99US-00232028.  
 PR 10-AUG-1999; 99US-00371347.  
 XX  
 PA (UIMC-) UNIV MCGILL.  
 XX  
 PI Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;  
 WPI; 2000-466131/40.  
 XX  
 DR Mammalian methionine synthase reductase nucleic acid used for detecting  
 PT an increased risk of developing a neural tube defect, Down's Syndrome or  
 PT cardiovascular disease in a mammalian embryo or fetus.  
 XX  
 PS Claim 8; Page; 85pp; English.  
 XX  
 CC The present sequence represents a human methionine synthase reductase  
 CC (MTRR) DNA sequence, with a polymorphism comprising of a deletion of  
 CC nucleotides 1726-1728. Inhibitors of MTRR polypeptide and polynucleotide  
 CC are used for treating or preventing cancer, cardiovascular disease,  
 CC Down's Syndrome or neural tube defects in a subject. The cardiovascular  
 CC disease is premature coronary artery disease. The compounds are detected  
 CC by methods which screen for modulators of MTRR biological activity. MTRR  
 CC polypeptide or nucleic acid is examined for the presence of a  
 CC polymorphism in the parents or the embryo or foetus, and the information  
 CC used for detecting an increased risk of an embryo or foetus developing  
 CC cancer, cardiovascular disease, Down's Syndrome or neural tube defects.  
 CC note: the present sequence does not appear in the specification; it was  
 CC created using information provided  
 CC  
 XX  
 SQ Sequence 3256 BP; 943 A; 705 C; 662 G; 946 T; 0 U; 0 Other;  
 Query Match 81.1%; Score 1701; DB 3; Length 3256;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 2091; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

1 ATGAGGAGGTTTCTGTTACTATATGCTACACAGCAGGAGCAGGCAAAAGCCATCGCAGAA 60  
80 ATGAGGAGGTTTCTGTTACTATATGCTACACAGCAGGAGCAGGCAAAAGCCATCGCAGAA 139  
61 GAAATGTGAGCAAGCTGTGTACATGATTTTCTGCAAGATCTTCACTATATTAAGTAA 120  
140 GAAATGTGAGCAAGCTGTGTACATGATTTTCTGCAAGATCTTCACTATATTAAGTAA 199  
121 TCCGATTAAGTAACTTAATAAAACCGAAACAGCTCTCTTGTGTGTGTGTCTTACACAG 180  
200 TCCGATTAAGTAACTTAATAAAACCGAAACAGCTCTCTTGTGTGTGTGTCTTACACAG 259  
181 GGCACCGGAGACCCACCGACACAGCCCGCAAGTTGTTAAGGAATAACGAACCAACA 240  
260 GGCACCGGAGACCCACCGACACAGCCCGCAAGTTGTTAAGGAATAACGAACCAACA 319  
241 CTGCGCGGTGATTTCTTTGCTCACTGCGGTATGGGTACTGGGTCTCGGTGATTCAGAA 300  
320 CTGCGCGGTGATTTCTTTGCTCACTGCGGTATGGGTACTGGGTCTCGGTGATTCAGAA 379  
301 TACACCTACTTTTGGCAATGGGGGGAAGATTAATTGATTAACGACTTCAAGAGCTTGGAGCC 360  
380 TACACCTACTTTTGGCAATGGGGGGAAGATTAATTGATTAACGACTTCAAGAGCTTGGAGCC 439  
361 CGGCATTTCTATGACA CTGGAACATGACATGACCTGTAGAGTTAGAACTTGTGTGTAG 420  
440 CGGCATTTCTATGACA CTGGAACATGACATGACCTGTAGAGTTAGAACTTGTGTGTAG 499  
421 CGGTGATTTGCTGAGACTGCGCAGCCCTCAGAAAGCATTTTATGATCAAGCAGAGACA 480  
500 CGGTGATTTGCTGAGACTGCGCAGCCCTCAGAAAGCATTTTATGATCAAGCAGAGACA 559  
481 GAGGAGATTAAGTGGCGCATCTCCGCTGGCATCACTGCACTCTTGAAGACAGACTTGTG 540  
560 GAGGAGATTAAGTGGCGCATCTCCGCTGGCATCACTGCACTCTTGAAGACAGACTTGTG 619  
541 AAGTCAGAGCTGTACACATGATATCTCAAGTGAAGCTTCTGAGTTTCAATGATTCAGAA 600  
620 AAGTCAGAGCTGTACACATGATATCTCAAGTGAAGCTTCTGAGTTTCAATGATTCAGAA 679  
601 AGAAAGATTTCTGAGTTTGAAGCAAAATGACAGTAAACAGCAACCAATCCAAATGTTGTA 660  
680 AGAAAGATTTCTGAGTTTGAAGCAAAATGACAGTAAACAGCAACCAATCCAAATGTTGTA 739  
661 ATTGAAGACTTGAAGTCTCACTTACCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
740 ATTGAAGACTTGAAGTCTCACTTACCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 799  
721 AATATTCGAGTTTACCCCGAATATTTAAGATTAATCTGAGAGATCTTGTGCGCAG 780  
800 AATATTCGAGTTTACCCCGAATATTTAAGATTAATCTGAGAGATCTTGTGCGCAG 859  
781 GAGGAAAGCAAGTATCTGAGCTTCAAGCAGATCTCAAGTTTCAAGTCAAGTTTCAAG 840  
860 GAGGAAAGCAAGTATCTGAGCTTCAAGCAGATCTCAAGTTTCAAGTCAAGTTTCAAG 919  
841 GCAGTTCAACTTACTACGAATGATGCAATTAACCACTCTGCTGAGTAATGAGCAATT 900  
920 GCAGTTCAACTTACTACGAATGATGCAATTAACCACTCTGCTGAGTAATGAGCAATT 979  
901 TCAAAATACAGACTTTTCTATACAGCTGAGAGATGCTTCAAGCTGATCTGCTCAACAGT 960  
980 TCAAAATACAGACTTTTCTATACAGCTGAGAGATGCTTCAAGCTGATCTGCTCAACAGT 1039  
961 GATTCTGAGGTACAAAGCCTACTCAAAAGCTGAGCTTGAAGTAAATAAGAGCACTGC 1020  
1040 GATTCTGAGGTACAAAGCCTACTCAAAAGCTGAGCTTGAAGTAAATAAGAGCACTGC 1099  
1021 GTCTCTTTGAAAAATTAAGGCAACACAAAGAAAGAGACTTACTTCCCGCAGCATATA 1080  
1100 GTCTCTTTGAAAAATTAAGGCAACACAAAGAAAGAGACTTACTTCCCGCAGCATATA 1159  
1081 CTGCGGGAATGTTCTCTCAGTTCAATTTTACCTGAGTGTGTAATAATCCAGCAATTCCT 1140

1160 CCGCGGGAATGTTCTCTCCAGTTCAATTTTACCTGGTGTCTGAAATCCGAGCAATTCCT 1219  
1141 AAAAAAGCATTTTGTGAGACCTTGTGAGCTTATACAGAGACAGCTGTAATAAGCCAGG 1200  
1220 AAAAAAGCATTTTGTGAGACCTTGTGAGCTTATACAGAGACAGCTGTAATAAGCCAGG 1279  
1201 CTACAGAGCTGTGAGTAAACAGAGGAGAGCCGATTAATAGCCGCTTGTATAGAGATGCC 1260  
1280 CTACAGAGCTGTGAGTAAACAGAGGAGAGCCGATTAATAGCCGCTTGTATAGAGATGCC 1339  
1261 TGTGCTGTGTGTGATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTC 1320  
1340 TGTGCTGTGTGTGATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTC 1399  
1321 CTGCTGGAACATCTTCTTAACTTCAACCCAGACCATATTTGTGTGAGCTTCAAGTTTA 1380  
1400 CTGCTGGAACATCTTCTTAACTTCAACCCAGACCATATTTGTGTGAGCTTCAAGTTTA 1459  
1381 TTTCACCCGAGAAAGCTCAATTTTGTCTTCAACATTTGTGAAATTTCTGCTACTGCGACA 1440  
1460 TTTCACCCGAGAAAGCTCAATTTTGTCTTCAACATTTGTGAAATTTCTGCTACTGCGACA 1519  
1441 ACAGAGTTCTGCGAAGGAGATATGATACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
1520 ACAGAGTTCTGCGAAGGAGATATGATACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1579  
1501 CTTCAGCCAAACATCATCATCATCTCCATGAAAGACAGCGGAAAGCCCTGCTCTTAAGATA 1560  
1580 CTTCAGCCAAACATCATCATCTCCATGAAAGACAGCGGAAAGCCCTGCTCTCTTAAGATA 1639  
1561 TCCATCTCTCTGCAACCAAAATCTTCCATTAACAGATGAGCCCTCAATCCCATC 1620  
1640 TCCATCTCTCTGCAACCAAAATCTTCCATTAACAGATGAGCCCTCAATCCCATC 1699  
1621 ATATGATGGGTCTGAGAAACCGGCAATGCCCCGTTATTTGTTTCTTCAACATAGAGAG 1680  
1700 ATATGATGGGTCTGAGAAACCGGCAATGCCCCGTTATTTGTTTCTTCAACATAGAGAG 1756  
1681 AAATCTCAAGAACCAACCCAGATGGAATTTTGAAGCATGTGTGTTTTTGTGCTGC 1740  
1757 AAATCTCAAGAACCAACCCAGATGGAATTTTGAAGCATGTGTGTTTTTGTGCTGC 1816  
1741 AGGCATTAAGATGAGATTAATCTATCAAGAAAGCTGAGCATTTCTTAAAGCATGGG 1800  
1817 AGGCATTAAGATGAGATTAATCTATCAAGAAAGCTGAGCATTTCTTAAAGCATGGG 1876  
1801 ATCTTAATCATCTAAAGGTTTCTTCTCAAGAGATGCTCTGTTGGGAGAGAGAGGCC 1860  
1877 ATCTTAATCATCTAAAGGTTTCTTCTCAAGAGATGCTCTGTTGGGAGAGAGAGGCC 1936  
1861 CCAGCAAGATATGATCAAGACCAATCATGAGCTTCAATGCGCAGAGAGTGGCAGATCTTC 1920  
1937 CCAGCAAGATATGATCAAGACCAATCATGAGCTTCAATGCGCAGAGAGTGGCAGATCTTC 1996  
1921 CTTCAGAGAAAGCGGCAATTTATATGTGTGTGAGATGCAAGAAATATGCGCAAGATGTA 1980  
1997 CTTCAGAGAAAGCGGCAATTTATATGTGTGTGAGATGCAAGAAATATGCGCAAGATGTA 2056  
1981 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGTTGAGTTGAAAACTAAGAGCAATG 2040  
2057 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGTTGAGTTGAAAACTAAGAGCAATG 2116  
2041 AAAACCTGCGCACTTTAAAGAGAAACGCTACCTTCAGATATTTGTGCTATA 2097  
2117 AAAACCTGCGCACTTTAAAGAGAAACGCTACCTTCAGATATTTGTGCTATA 2173

RESULT 12  
AAAS8976  
ID AAAS8976 standard; DNA; 3255 BP.  
XX  
AC AAAS8976;



XX 07-NOV-2000 (first entry)  
XX A human methionine synthase reductase DNA sequence with polymorphism.  
XX  
XX  
XX  
XX  
XX Human: methionine synthase reductase; MTRR; cancer;  
XX cardiovascular disease; Down's Syndrome; neural tube defect;  
XX prematute coronary artery disease; ss.  
XX Homo sapiens.  
XX  
XX W0200042196-A2.  
XX  
XX 20-JUL-2000.  
XX  
XX 14-JAN-2000; 2000MO-IB000209.  
XX  
XX 15-JAN-1999; 99US-00232028.  
XX 10-AUG-1999; 99US-00371347.  
XX (UYMC-) UNIV MCGILL.  
XX  
XX Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;  
XX WPI; 2000-466131/40.  
XX  
XX Mammalian methionine synthase reductase nucleic acid used for detecting  
XX an increased risk of developing a neural tube defect. Down's Syndrome or  
XX cardiovascular disease in a mammalian embryo or fetus.  
XX  
XX Claim 7; Page; 85pp; English.  
XX  
XX The present sequence represents a human methionine synthase reductase  
XX (MTRR) DNA sequence, with a polymorphism comprising of a deletion of  
XX nucleotides 1675-1678. Inhibitors of MTRR polypeptide and polynucleotide  
XX are used for treating or preventing cancer, cardiovascular disease,  
XX Down's Syndrome or neural tube defects in a subject. The cardiovascular  
XX disease is prematute coronary artery disease. The compounds are detected  
XX by methods which screen for modulators of MTRR biological activity. MTRR  
XX polypeptide or nucleic acid is examined for the presence of a  
XX polymorphism in the parents or the embryo or foetus, and the information  
XX used for detecting an increased risk of an embryo or foetus developing  
XX cancer, cardiovascular disease, Down's Syndrome or neural tube defects.  
XX note: the present sequence does not appear in the specification; it was  
XX created using information provided  
XX  
XX Sequence 3255 BP; 942 A; 704 C; 663 G; 946 T; 0 U; 0 Other;  
XX  
XX Query Match 78.2%; Score 1640; DB 3; Length 3255;  
XX Best Local Similarity 99.7%; Pred. No. 0;  
XX Matches 2090; Conservative 0; Mismatches 3; Indels 4; Gaps 1;  
XX  
XX 1 ATGAGAGAGTTTCTGTACTATATGCTACACAGAGGAGGACGAGGAGCATGCGAGAA 60  
XX |||||||  
XX 80 ATGAGAGAGTTTCTGTACTATATGCTACACAGAGGAGGAGGAGGAGGAGGAGGAA 139  
XX |||||||  
XX 61 GAAATGTGTGAGCAAGCTGTGTGATCATGATTTTCTGCAATCTTCACTATATTAGTAA 120  
XX |||||||  
XX 140 GAAATGTGTGAGCAAGCTGTGTGATCATGATTTTCTGCAATCTTCACTATATTAGTAA 139  
XX |||||||  
XX 121 TCCGATTAAGTATGACTTAAACCGGAAACGAGCTCTCTGTGTGTGTGTGTGTCTACACG 180  
XX |||||||  
XX 200 TCCATTAAGTATGACTTAAACCGGAAACGAGCTCTCTGTGTGTGTGTGTGTCTACACG 259  
XX |||||||  
XX 181 GGCACCGGAGACCCAGCCGACACAGCCGCGAAGTTTGTAAAGAAATACAGAACCAACA 240  
XX |||||||  
XX 260 GGCACCGGAGACCCAGCCGACACAGCCGCGAAGTTTGTAAAGAAATACAGAACCAACA 319  
XX |||||||  
XX 241 CTGCGGTTGATTTCTTTGTCTCACTGCGGTATGGGTATCGGTCTCGGTATTCAGAA 300  
XX |||||||  
XX 320 CTGCGGTTGATTTCTTTGTCTCACTGCGGTATGGGTATCGGTCTCGGTATTCAGAA 379  
XX |||||||  
XX 301 TACACCTACTTTTGCAATGGGGGGAAGATTAATGATTAAGCACTTCAAGAGCTTGGAGCC 360

DB 380 TACACCTACTTTTGCAATGGGGGGAAGATTAATGATTAAGCACTTCAAGAGCTTGGAGCC 439  
|||  
QY CGGATTTTCTATGACACTGACATGACAGATGACTGTGAGTTTAAAGCTTGTGTTGAG 420  
|||  
DB CGGATTTTCTATGACACTGACATGACAGATGACTGTGAGTTTAAAGCTTGTGTTGAG 499  
|||  
QY CCGTGATTTGCTGACTGTGCGCCAGCCCTCAGAAACATTTTAAAGTCAAGCAGAGACA 480  
|||  
DB CCGTGATTTGCTGACTGTGCGCCAGCCCTCAGAAACATTTTAAAGTCAAGCAGAGACA 559  
|||  
QY GAGGATTAAGTGGCGCACTCCCGTGGCATACCTGCAATCTTGAAGGACAGACCTGTG 540  
|||  
DB GAGGATTAAGTGGCGCACTCCCGTGGCATACCTGCAATCTTGAAGGACAGACCTGTG 619  
|||  
QY AAGTCAGAGCTGCTACATTTGATATCTCAAGTCAAGCTTCTGAGATTGCAATTCAGGA 600  
|||  
DB AAGTCAGAGCTGCTACATTTGATATCTCAAGTCAAGCTTCTGAGATTGCAATTCAGGA 679  
|||  
QY AGAAGGATTTGAGGTTTGAAGCAAAATGCAGTGAACAGCAACCAATTCATGTTGTA 660  
|||  
DB AGAAGGATTTGAGGTTTGAAGCAAAATGCAGTGAACAGCAACCAATTCATGTTGTA 739  
|||  
QY ATTGAAGCTTTGAGTCTCACTTACCCGTTGGTACCCCACTCTCAAGAGCTGCTG 720  
|||  
DB ATTGAAGCTTTGAGTCTCACTTACCCGTTGGTACCCCACTCTCAAGAGCTGCTG 799  
|||  
QY AATATTCCTGTTTAAACCCAGAAATTTTAAAGTCAATCTGACAGAGTCTTGGCCAG 780  
|||  
DB AATATTCCTGTTTAAACCCAGAAATTTTAAAGTCAATCTGACAGAGTCTTGGCCAG 859  
|||  
QY 781 GAGGAAAGCCAGATATCTGTGACTTACAGAGATCCAGTTTAAAGTCCAAATTCAGAG 840  
|||  
DB GAGGAAAGCCAGATATCTGTGACTTACAGAGATCCAGTTTAAAGTCCAAATTCAGAG 919  
|||  
QY 841 GCGATTCATCTTAAGATATGCAATGATGCAATGCAATGCAATGCAATGCAATGCAAT 900  
|||  
DB GCGATTCATCTTAAGATATGCAATGATGCAATGCAATGCAATGCAATGCAATGCAAT 979  
|||  
QY 901 TCAATATCAAGACTTTCTATCAAGCTGAGAGTCTTCAAGCTGATCTGCTTCAAGT 960  
|||  
DB TCAATATCAAGACTTTCTATCAAGCTGAGAGTCTTCAAGCTGATCTGCTTCAAGT 979  
|||  
QY 980 TCAATATCAAGACTTTCTATCAAGCTGAGAGTCTTCAAGCTGATCTGCTTCAAGT 1039  
|||  
DB 961 GATTCTGAGTCAAGAGCTTCTCAAGAGCTGAGAGTCTGAGAGTCTGAGAGTCTGAG 1020  
|||  
QY 1040 GATTCTGAGTCAAGAGCTTCTCAAGAGCTGAGAGTCTGAGAGTCTGAGAGTCTGAG 1099  
|||  
DB 1021 GTCCTTTGAAATTAAGGAGACACAAAGAAAGAGGCTTACCCGAGCATATA 1080  
|||  
QY 1100 GTCCTTTGAAATTAAGGAGACACAAAGAAAGAGGCTTACCCGAGCATATA 1159  
|||  
DB 1081 CCGCGGAGATGTTCTCTCCAGTTCAATTTTAACTGTGTCTGAAATCCGAGCAATTCCT 1140  
|||  
QY 1160 CCGCGGAGATGTTCTCTCCAGTTCAATTTTAACTGTGTCTGAAATCCGAGCAATTCCT 1219  
|||  
DB 1141 AAAAAGCAATTTTGGAGAGCCCTTGTGACTATACAGTGAACGTGTGAAAGCCGAGG 1200  
|||  
QY 1220 AAAAAGCAATTTTGGAGAGCCCTTGTGACTATACAGTGAACGTGTGAAAGCCGAGG 1279  
|||  
DB 1201 CTACAGAGCTGTGAGTAAACAAAGGAGGAGCCGATTAATAGCCGCTTGTGAGAGATGCC 1260  
|||  
QY 1280 CTACAGAGCTGTGAGTAAACAAAGGAGGAGCCGATTAATAGCCGCTTGTGAGAGATGCC 1339  
|||  
DB 1261 TGTGCTGTGTTGATCTCTCTCGTTTCCCTTCTTCCAGGACCACTCACTCTC 1320  
|||  
QY 1340 TGTGCTGTGTTGATCTCTCTCGTTTCCCTTCTTCCAGGACCACTCACTCTC 1399  
|||  
DB 1321 CTGCTGGAACATCTTCTTAACTTCAACCAAGCAATATTCGTGCAAGCTCAAGTTTA 1380  
|||  
QY 1400 CTGCTGGAACATCTTCTTAACTTCAACCAAGCAATATTCGTGCAAGCTCAAGTTTA 1459  
|||  
DB 1381 TTTCACCCAGGAAGCTTCAATTTTGTCTTCAATATGTGGAATTTTGTCTTACCTGACA 1440  
|||



Db 1460 TTTCACCCGAGAAAGCTCATTGTTCTTCAATGTTGAAATTTCTGTCTACTGCCACA 1519  
 Qy 1441 ACAGAGTTCTGCGGAGGAGATGTATACAGGCTGGCTGCGCTTGTTGGTCTTCAGTT 1500  
 Db 1520 ACAGAGTTCTGCGGAGGAGATGTATACAGGCTGGCTGCGCTTGTTGGTCTTCAGTT 1579  
 Qy 1501 CTTCAGCCAAACATACATGATCCCATGAAAGACAGCGGAGAAAGCCCTGGCTCTAAGATA 1560  
 Db 1580 CTTCAGCCAAACATACATGATCCCATGAAAGACAGCGGAGAAAGCCCTGGCTCTAAGATA 1639  
 Qy 1561 TCCATCTCTCTCGAACAACAATTTCTTCACCTTACCAAGATGACCCCTCAATCCCATC 1620  
 Db 1640 TCCATCTCTCTCGAACAACAATTTCTTCACCTTACCAAGATGACCCCTCAATCCCATC 1695  
 Qy 1621 ATAAATGTTGGTCCAGAACCGGCAATAGCCCGTTATTTGGTCTTACCAATGAGAG 1680  
 Db 1696 ATAAATGTTGGTCCAGAACCGGCAATAGCCCGTTATTTGGTCTTACCAATGAGAG 1755  
 Qy 1681 AAATCTCAGAACAAACCCAGATGGAATTTTGGAGCAATGTGTTGTTTTTGGCTGC 1740  
 Db 1756 AAATCTCAGAACAAACCCAGATGGAATTTTGGAGCAATGTGTTGTTTTTGGCTGC 1815  
 Qy 1741 AGGCATTAAGATAGGATTAATCTTATTCAGAAAAAGCTCAGACATTTCTTAAGCATGCG 1800  
 Db 1816 AGGCATTAAGATAGGATTAATCTTATTCAGAAAAAGCTCAGACATTTCTTAAGCATGCG 1875  
 Qy 1801 ATCTTAACATCATTAAGTTTCTTCTCAAGAGATGCTCCTGTTGGAGAGAGAGAGCC 1860  
 Db 1876 ATCTTAACATCATTAAGTTTCTTCTCAAGAGATGCTCCTGTTGGAGAGAGAGAGCC 1935  
 Qy 1861 CCAGCAAAATATGTACAAAGACATCCAGCTTCATGCGCAGCAGGTCGAGAAATCTCTC 1920  
 Db 1936 CCAGCAAAATATGTACAAAGACATCCAGCTTCATGCGCAGCAGGTCGAGAAATCTCTC 1995  
 Qy 1921 CTCACAGAGACCGCCCATATTTATGTGTGTGAGATGACAAAGATATGSCCAAGATGTA 1980  
 Db 1996 CTCACAGAGACCGCCCATATTTATGTGTGTGAGATGACAAAGATATGSCCAAGATGTA 2055  
 Qy 1981 CATGATGCCCTTGTGCAATTAATAGCAAGAGTTGAGTTGAAAAACTTAAGCAATG 2040  
 Db 2056 CATGATGCCCTTGTGCAATTAATAGCAAGAGTTGAGTTGAAAAACTTAAGCAATG 2115  
 Qy 2041 AAAACCTGGCCCACTTTAAAGAAAGAAAGCTACCTTCAGAGATTTTGGTCATTA 2097  
 Db 2116 AAAACCTGGCCCACTTTAAAGAAAGAAAGCTACCTTCAGAGATTTTGGTCATTA 2172

## RESULT 13

ADQ39029  
ID ADQ39029 standard; DNA; 3256 BP.

XX ADQ39029;

XX 18-NOV-2004 (first entry)

XX Human SNP containing myocardial infarction-associated gene, SEQ ID 692.

XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;  
 KW cardiant; gene therapy; human; gene; ds.

OS Homo sapiens.

XX PN WO2004058052-A2.

XX PD 15-JUL-2004.

XX PF 22-DEC-2003; 2003WO-US040978.

XX PR 20-DEC-2002; 2002US-0434778P.

XX PR 10-MAR-2003; 2003US-0453135P.

XX PR 30-APR-2003; 2003US-0466412P.

XX PR 23-SEP-2003; 2003US-0504955P.

PA (APPL-) APPLERA CORP.  
 XX  
 PI Cargill M, Devlin J, Iakoubova O;  
 XX  
 XX WPI, 2004-533949/51.  
 DR P-PSDB; ADQ39857.  
 DR  
 XX  
 XX Identifying an individual who has an altered risk for developing  
 PT myocardial infarction by detecting a single nucleotide polymorphism in  
 PT the individual's nucleic acids.  
 XX  
 XX  
 PS Claim 7; SEQ ID NO 692; 145bp; English.  
 CC  
 CC The invention relates to a novel method for identifying an individual who  
 CC has an altered risk for developing myocardial infarction. The method  
 CC comprises detecting a single nucleotide polymorphism (SNP) in any one of  
 CC the nucleotide sequences given in the specification in the individual's  
 CC nucleic acids, where the presence of the SNP is correlated with an  
 CC altered risk for myocardial infarction in the individual. The invention  
 CC further comprises: an isolated nucleic acid molecule comprising at least  
 CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in  
 CC the specification or its complement and encoding any one of the amino  
 CC acid sequences given in the specification; an isolated polypeptide  
 CC comprising an amino acid sequence given in the specification; an antibody  
 CC that specifically binds to the polypeptide or its antigen-binding  
 CC fragment; an amplified polynucleotide containing an SNP given in the  
 CC specification and which is between about 16 and 1000 nucleotides in  
 CC length; a kit for detecting an SNP in a nucleic acid, comprising the  
 CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a  
 CC nucleic acid molecule; a method of detecting a variant polypeptide; and a  
 CC method for identifying an agent useful in treating or preventing  
 CC myocardial infarction. The novel detection method has cardiant activity.  
 CC The nucleic acids of the invention may be used in gene therapy. The  
 CC method is useful in identifying an individual who has an increased or  
 CC decreased risk for developing myocardial infarction and for preparing a  
 CC composition for treating or preventing myocardial infarction. This  
 CC polynucleotide sequence represents a human myocardial infarction-  
 CC associated gene containing one or more SNP's of the invention. Note: This  
 CC sequence was not shown in the specification. The sequence has come from  
 CC an electronic sequence listing downloaded from the WIPO website.  
 XX  
 SQ Sequence 3256 BP; 927 A; 691 C; 669 G; 940 T; 0 U; 29 Other;  
 Query Match 48.5%; Score 1018; DB 13; Length 3256;  
 Best Local Similarity 99.0%; Pred. No. 0;  
 Matches 1968; Conservative 0; Mismatches 19; Indels 0; Gaps 0;  
 Qy 111 TATTAGTGAATCCGATTAAGTATGACTTAAMAAACGAAACAGCTCCTGTTGTTGTGTT 170  
 Db 204 TATTAGTGAATCCGATTAAGTATGACTTAAMAAACGAAACAGCTCCTGTTGTTGTGTT 263  
 Qy 171 TTCTACACAGGCGACCGGAGACCCACCGACACAGCCCGCAAGTTTGTAAAGAAATACA 230  
 Db 264 TTCTACACAGGCGACCGGAGACCCACCGACACAGCCCGCAAGTTTGTAAAGAAATACA 323  
 Qy 231 GAACCAACACGCGCGTGAATTTCTTGTCTACCGTGGGATATGAGTTAGCTGGCTGG 290  
 Db 324 GAACCAACACGCGCGTGAATTTCTTGTCTACCGTGGGATATGAGTTAGCTGGCTGG 383  
 Qy 291 TGAATTCAGATACACCTACTTTTGCATATGAGGAGAAATATGATTAACGACTTCAAGA 350  
 Db 384 TGAATTCAGATACACCTACTTTTGCATATGAGGAGAAATATGATTAACGACTTCAAGA 443  
 Qy 351 GCTTGAAGCCCGCAATTTCTATGACATGAGACATGACATGCTGTGATGTTAGAACT 410  
 Db 444 GCTTGAAGCCCGCAATTTCTATGACATGAGACATGACATGCTGTGATGTTAGAACT 503  
 Qy 411 TGTGTTGAGCGGTGATGCTGAGCTGCGCACCGCTCAGAAAGCATTTAGGTCAG 470  
 Db 504 TGTGTTGAGCGGTGATGCTGAGCTGCGCACCGCTCAGAAAGCATTTAGGTCAG 563  
 Qy 471 CAGAGACAAAGAGATTAAGTGGCGCACTCCCGGTGGCATGCTGCAATCCTTGAAGAC 530

D 564 CAGAGCAAGAGAGATAGTGGCCGACCTCCGGTGGATCACTGATCTCTTGAGGAC 623  
Q 531 AGACTTTGTAGAGTCAAGTCTCTACACATGTAATCTCAAGTCAAGTCTTGAGATTG 590  
D 624 AGACCTGTGTAAGTCAAGTCTCTACACATGTAATCTCAAGTCAAGTCTTGAGATTG 683  
Q 591 TGAATTCAGAGAAAGAGATCTGAGTTTGAAGCAAAATGCAATGCAACGCAACCAATC 650  
D 684 TGAATTCAGAGAAAGAGATCTGAGTTTGAAGCAAAATGCAATGCAACGCAACCAATC 743  
Q 651 CAATGTTGTAATGAAAGACTTTGAGTCTCACTTACCCGTTGCGTACCCCACTCTACA 710  
D 744 CAATGTTGTAATGAAAGACTTTGAGTCTCACTTACCCGTTGCGTACCCCACTCTACA 803  
Q 711 AGCTCTCTGAATATCTCTGTTTACCCGAAATATTTACAGTACATCTGACAGAGTGC 770  
D 804 AGCTCTCTGAATATCTCTGTTTACCCGAAATATTTACAGTACATCTGACAGAGTGC 863  
Q 771 TCTTGGCAGAGAAAGCAAGATCTGAGTCTTCAAGATCTTCAAGTCTTCAAGTCTG 830  
D 864 TCTTGGCAGAGAAAGCAAGATCTGAGTCTTCAAGATCTTCAAGTCTTCAAGTCTG 923  
Q 831 AATTTCAAGAGAGTCACTTACTAGATGATGCAATGATGCAATGATGCAATGATG 890  
D 924 AATTTCAAGAGAGTCACTTACTAGATGATGCAATGATGCAATGATGCAATGATG 983  
Q 891 ATTGGAATTTCAATACAGACTTTTCTATCAGCTGAGATGCTTCAAGTCTGATCTG 950  
D 984 ATTGGAATTTCAATACAGACTTTTCTATCAGCTGAGATGCTTCAAGTCTGATCTG 1043  
Q 951 CCTTACAGATGATCTGAGTCAAGAGTCTTCAAGATGCTGAGTCTTCAAGTCTG 1010  
D 1044 CCTTACAGATGATCTGAGTCAAGAGTCTTCAAGATGCTGAGTCTTCAAGTCTG 1103  
Q 1011 AGAGCACTGCTCTTTGAAATTAAGCAGACACAAGAAAGAAAGAGTCACTTAC 1070  
D 1104 AGAGCACTGCTCTTTGAAATTAAGCAGACACAAGAAAGAAAGAGTCACTTAC 1163  
Q 1071 CAGACATATACCTGCGGAGTCTCTCAAGTCAATTTTACCTGATGCTTGAATCCG 1130  
D 1164 CAGACATATACCTGCGGAGTCTCTCAAGTCAATTTTACCTGATGCTTGAATCCG 1223  
Q 1131 AGCAATTTCTTAAAAAGCATTTTGGCAGCCCTTGTGATCTTACCGATGATGCTG 1190  
D 1224 AGCAATTTCTTAAAAAGCATTTTGGCAGCCCTTGTGATCTTACCGATGATGCTG 1283  
Q 1191 AAAGCGAGGCTACAGAGTCTGTCAGTAAACAAAGGCGCAGCCATATAGCCCTTGT 1250  
D 1284 AAAGCGAGGCTACAGAGTCTGTCAGTAAACAAAGGCGCAGCCATATAGCCCTTGT 1343  
Q 1251 ACGAGATGCTGCTGCTGCTTGTGATCTCTCTGCTTCTTCCCTTCTTCCAGCCACC 1310  
D 1344 ACGAGATGCTGCTGCTGCTTGTGATCTCTCTGCTTCTTCCCTTCTTCCAGCCACC 1403  
Q 1311 ACTCAGTCTCTGCTGCTGCTTGTGATCTCTCTGCTTCTTCCCTTCTTCCAGCCACC 1370  
D 1404 ACTCAGTCTCTGCTGCTGCTTGTGATCTCTCTGCTTCTTCCCTTCTTCCAGCCACC 1463  
Q 1371 CTCAGATTTATTTCAACCGAAAGCTCAATTTGTCTTCAACCTTGTGAAATTTCTGTC 1430  
D 1464 CTCAGATTTATTTCAACCGAAAGCTCAATTTGTCTTCAACCTTGTGAAATTTCTGTC 1523  
Q 1431 TACTGCAACAACAGAGTCTTGCAGAAAGAGATGTAACAGGCTGCTGCTTGTGAT 1490  
D 1524 TACTGCAACAACAGAGTCTTGCAGAAAGAGATGTAACAGGCTGCTGCTTGTGAT 1583  
Q 1491 TGCTTCAGTCTTCAAGCCAAACATACATGATCTCCATGAAAGACGCGGAAAGCCCTG 1550  
D 1584 TGCTTCAGTCTTCAAGCCAAACATACATGATCTCCATGAAAGACGCGGAAAGCCCTG 1643  
Q 1551 TCCTAAGATATTCATCTCTCTCGAACAACAAATTTCTTCACTTACAGATGACCCCTC 1610  
D 1644 TCCTAAGATATTCATCTCTCTCGAACAACAAATTTCTTCACTTACAGATGACCCCTC 1703

Q 1611 AATCCCATCATATATGTTGGTTCAGAAACCGGATAGCCCGTTTATTTGGTCTCTACA 1670  
D 1704 AATCCCATCATATATGTTGGTTCAGAAACCGGATAGCCCGTTTATTTGGTCTCTACA 1763  
Q 1671 ACATAGAGAAATCTCCAGAAACAACCCAGATGAAATTTTGGCAATGTTGTTT 1730  
D 1764 ACATAGAGAAATCTCCAGAAACAACCCAGATGAAATTTTGGCAATGTTGTTT 1823  
Q 1731 TTTTGGCTCAGGCAATAGATAGGATATATCTATTCAGAAAAAGCTCAGATTTCT 1790  
D 1824 TTTTGGCTCAGGCAATAGATAGGATATATCTATTCAGAAAAAGCTCAGATTTCT 1883  
Q 1791 TAAAGATGGGATCTTAACTCATCTTAAAGTTTCTTCTCAAGAGTCTCTGTTGGGA 1850  
D 1884 TAAAGATGGGATCTTAACTCATCTTAAAGTTTCTTCTCAAGAGTCTCTGTTGGGA 1943  
Q 1851 GAGAGAGCCCGCAGCAAGATATGTAACAACAACATCCAGTTCATGAGCAGAGTGGC 1910  
D 1944 GAGAGAGCCCGCAGCAAGATATGTAACAACAACATCCAGTTCATGAGCAGAGTGGC 2003  
Q 1911 GAGATCTCTCTCAGAGAAAGCCCATATTTATGTTGTGAGATGCAAGATATG 1970  
D 2004 GAGATCTCTCTCAGAGAAAGCCCATATTTATGTTGTGAGATGCAAGATATG 2063  
Q 1971 CAAAGATGTAATGATGCTCTTGTGCAATTAATAGCAAAAGCTTGGAGTTGAAA 2030  
D 2064 CAAAGATGTAATGATGCTCTTGTGCAATTAATAGCAAAAGCTTGGAGTTGAAA 2123  
Q 2031 AGAAGCAATGAAAGCTGCGCCTTAAAGAAAGAAAGCTTCAAGATATTTG 2090  
D 2124 AGAAGCAATGAAAGCTGCGCCTTAAAGAAAGAAAGCTTCAAGATATTTG 2183  
Q 2091 GTCATTA 2097  
D 2184 GTCATTA 2190

RESULT 14  
ADQ39030  
ID ADQ39030 standard; DNA; 3274 BP.  
XX  
AC ADQ39030;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Human SNP containing myocardial infarction-associated gene, SEQ ID 693.  
XX  
DE Myocardial infarction; detection; single nucleotide polymorphism; SNP;  
KW cardiant; gene therapy; human; gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN W02004058052-A2.  
XX  
PD 15-JUL-2004.  
XX  
PF 22-DEC-2003; 2003WO-US040978.  
XX  
PR 20-DEC-2002; 2002US-0434778P.  
PR 10-MAR-2003; 2003US-0453135P.  
PR 30-APR-2003; 2003US-046412P.  
PR 23-SEP-2003; 2003US-0504955P.  
XX  
PA (APPL-) APPLERA CORP.  
XX  
PI Gargill M, Devlin JT, Iakoubova O;  
XX  
XX WPI; 2004-533949/51.  
DR P-PSDB; ADQ39858.  
XX  
PT Identifying an individual who has an altered risk for developing  
PT myocardial infarction by detecting a single nucleotide polymorphism in



QY 1731 TTTTGGCTGAGGATAGAGATAGAGATATCTATTTCAGAAAAGAGCTCAGACATTTCTT 1790  
DB 1842 TTTTGGCTGAGGATAGAGATAGAGATATCTATTTCAGAAAAGAGCTCAGAAATTTCTT 1901  
QY 1791 TAAGCATGGATCTTAATCATCTAAAGTTTCTCTCAAGAGATCTCTGTGGGGA 1850  
DB 1902 TAAGCATGGATCTTAATCATCTAAAGTTTCTCTCAAGAGATCTCTGTGGGGA 1961  
QY 1851 GAGAGAAAGCCCGAGCAAAAGTATGTACAGACAAATCAGCTTCATGGCCAGAGTGGC 1910  
DB 1962 GAGAGAAAGCCCGAGCAAAAGTATGTACAGACAAATCAGCTTCATGGCCAGAGTGGC 2021  
QY 1911 GAGATCTCTCTCAGAGAGAGAGGCGATTTTATGTGTGGAGATGCAAAAGATTTGCG 1970  
DB 2022 TAGATCTCTCTCAGAGAGAGAGGCGATTTTATGTGTGGAGATGCAAAAGATTTGCG 2081  
QY 1971 CAAGATGTATCATGATGCTCTGTGTGCAAAATTAATAGCAAAAGAGTTGAGTTGAAAAC 2030  
DB 2082 CAAGATGTATCATGATGCTCTGTGTGCAAAATTAATAGCAAAAGAGTTGAGTTGAAAAC 2141  
QY 2031 AGAAGCAATGAAAACCTTGCGCACTTTAAAGAAAAGAAACGCTACCTTCAGAGATTTTG 2090  
DB 2142 AGAAGCAATGAAAACCTTGCGCACTTTAAAGAAAAGAAACGCTACCTTCAGAGATTTTG 2201  
QY 2091 GTCATTA 2097  
DB 2202 GTCATTA 2208

RESULT 15  
ID ACN42470 standard; cDNA; 3189 BP.  
AC ACN42470;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Human diagnostic and therapeutic polynucleotide SEQ ID NO:1345.  
XX  
KW se; gene; gene therapy; human diagnostic and therapeutic polynucleotide;  
XX dithp.  
XX  
OS Homo sapiens.  
XX  
PN MO2004023973-A2.  
PD 25-MAR-2004.  
XX  
PE 12-SEP-2003; 2003MO-US028227.  
XX  
PR 12-SEP-2002; 2002US-0410259P.  
XX 12-SEP-2002; 2002US-0410260P.  
XX  
PA (INCY-) INCYTE CORP.  
XX  
PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;  
PI Hartshorne TA, Suchorolski MT, Altus CM, Pites SD, Elder LV;  
PI Mooney EM, Deleage AM, Panesar IS, Barville SC, Reddy TP;  
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstein EH;  
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LI;  
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Viltz UA, Kitron BS;  
PI Xu Y, Kong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;  
PI Patry S, Shi X, Suarez CJ;  
XX  
XX WPI; 2004-329368/30.  
DR P-PSDB; ABM83818.  
XX  
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful  
PT in diagnosing a condition, disease or disorder associated with human  
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or  
PT in gene mapping.  
PS Claim 1; Page; 190pp; English.

XX The invention relates to novel diagnostic and therapeutic polynucleotides  
CC selected from one of the 2722 sequences defined in the specification. A  
CC polynucleotide of the invention may have a use in gene therapy. The human  
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be  
CC used to diagnose a particular condition, disease or disorder associated  
CC with human molecules, e.g. cell proliferative disorder,  
CC autoimmune/inflammatory disorder, developmental disorder, endocrine  
CC disorder, neurological disorders, gastrointestinal disorders, or  
CC infections caused by virus, bacteria, fungi or parasite. The dithp  
CC molecules may also be used in genetic mapping, in identifying individuals  
CC from minute biological samples, in detecting single nucleotide  
CC polymorphisms, as molecular weight markers, and for somatic or germ-line  
CC gene therapy. The present sequence represents a dithp polynucleotide of  
CC the invention. Note: The sequence data for this patent is not represented  
CC in the printed specification, but was obtained in electronic format  
CC directly from WIPO at [www.wipo.int/pct/en/sequences/listing.htm](http://www.wipo.int/pct/en/sequences/listing.htm)  
XX  
SQ Sequence 3189 BP; 916 A; 679 C; 665 G; 929 T; 0 U; 0 Other;  
Query Match 43.2%; Score 905; DB 13; Length 3189;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1055; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 ANAGAGAGTTTCTGTTACTATATGCTATACAGACAGAGGACAGGCAAGGCGATCGAGAA 60  
DB 112 ATGAGAGAGTTTCTGTTACTATATGCTATACAGACAGAGGACAGGCAAGGCGATCGAGAA 171  
QY 61 GAAATGTGAGCAAGCTGTGTACATGATTTTCTGAGATCTTCAATATATATGAA 120  
DB 172 GAAATGTGAGCAAGCTGTGTACATGATTTTCTGAGATCTTCAATATATGAA 221  
QY 121 TCCGATAGTATGACCTTAAACCGAAGACGCTCTTGTGTGTGTTTCTACACG 180  
DB 232 TCCGATAGTATGACCTTAAACCGAAGACGCTCTTGTGTGTGTTTCTACACG 291  
QY 181 GGCACCGGAGACCCACCGACACAGCCCGCAAGTTTGAAGAAATACAGAACCAAC 240  
DB 292 GGCACCGGAGACCCACCGACACAGCCCGCAAGTTTGAAGAAATACAGAACCAAC 351  
QY 241 CTGCGGTTGATTTCTTGTCTCACTGCGGTATGAGTTTCTGAGTCTCGGTATTCAGAA 300  
DB 352 CTGCGGTTGATTTCTTGTCTCACTGCGGTATGAGTTTCTGAGTCTCGGTATTCAGAA 411  
QY 301 TACACCTACTTTTGCAGTGGGGGAAATATATGATTAACGACTTCAAGGCTTGAAGCC 360  
DB 412 TACACCTACTTTTGCAGTGGGGGAAATATGATTAACGACTTCAAGGCTTGAAGCC 471  
QY 361 CGGCAATTTCTATGACACTGACATGAGATGACTGTGATTAACCTTGTGTTGAG 420  
DB 472 CGGCAATTTCTATGACACTGACATGAGATGACTGTGATTAACCTTGTGTTGAG 531  
QY 421 CGGTGATTTCTGAGCTCTGCGCAGCGCTCAGAAACATTTTATGATCAAGAGAGCAA 480  
DB 532 CGGTGATTTCTGAGCTCTGCGCAGCGCTCAGAAACATTTTATGATCAAGAGAGCAA 591  
QY 481 GAGAGATTAAGTGGCGCACTCCGTTGGCATCCTGCACTTGAAGACAGACTTTG 540  
DB 592 GAGAGATTAAGTGGCGCACTCCGTTGGCATCCTGCACTTGAAGACAGACTTTG 651  
QY 541 AATCTAGACTGTATACATGAAATCTCAAGTGCAGCTTCTGAGATTCAGATTTGAGGA 600  
DB 652 AATCTAGACTGTATACATGAAATCTCAAGTGCAGCTTCTGAGATTCAGATTTGAGGA 711  
QY 601 AGAAGATTTCTGAGGTTTGAAGCAAAATGACAGTGAACAGCAACCAATCCATGTTGTA 660  
DB 712 AGAAGATTTCTGAGGTTTGAAGCAAAATGACAGTGAACAGCAACCAATCCATGTTGTA 771  
QY 661 ATTGAAGACTTTGAGTCTTCACTTACCCGTTGCTGATCCCTCACTTCAAGGCTTCTG 720  
DB 772 ATTGAAGACTTTGAGTCTTCACTTACCCGTTGCTGATCCCTCACTTCAAGGCTTCTG 831  
QY 721 AATTTCTGTTTACCCCAAGAAATTTTACAGGTATCTGACAGAGTCTCTGGCCAG 780

```
Db      832 AATATTCCTGTTTACCCCGAATATTACAGGTACATCTGCAGAGTCTCTGGCCAG 891
Qy      781 GAGGAAGCGAATATCTGTGACTTCAGAGATCCAGTTTCAAGTGCATTTCAAG 840
Db      892 GAGGAAGCGAATATCTGTGACTTCAGAGATCCAGTTTCAAGTGCATTTCAAG 951
Qy      841 GCAGTTCAACTTACTACGAATGATGSCCATAAAAACCACTCTGTGTAGATTGACATT 900
Db      952 GCAGTTCAACTTACTACGAATGATGSCCATAAAAACCACTCTGTGTAGATTGACATT 1011
Qy      901 TCAAAATACAGACTTTTCTATCAGCCTGAGATGCTTCAGCGTGAATCTGCCCTAACAGT 960
Db      1012 TCAAAATACAGACTTTTCTATCAGCCTGAGATGCTTCAGCGTGAATCTGCCCTAACAGT 1071
Qy      961 GATTCTGAGGTACAAAGCCTACTCCAAAGACTGCGAGCTTGAAAGATTAAGAGACACTGC 1020
Db      1072 GATTCTGAGGTACAAAGCCTACTCCAAAGACTGCGAGCTTGAAAGATTAAGAGACACTGC 1131
Qy      1021 GTCCCTTTGAAATAAGGACAGACAAAGAAAGAG 1058
Db      1132 GTCCCTTTGAAATAAGGACAGACAAAGAAAGAG 1169
```

Search completed: August 27, 2005, 01:18:36  
Job time : 738.49 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 23:30:25 ; Search time 235.757 Seconds  
(without alignments)

14554.251 Million cell updates/sec

Title: US-09-371-347a-43

Perfect score: 2097  
Sequence: 1 atgagagagcttcgtact.....ttcagacatttgcataa 2097

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq:\*
- 2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq:\*
- 3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq:\*
- 4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*
- 5: /cgn2\_6/prodata/1/ina/6C\_COMB.seq:\*
- 6: /cgn2\_6/prodata/1/ina/6D\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2046	97.6	3259	3	US-09-318-448-23 Sequence 23, Appl
2	1893	90.3	3242	4	US-09-949-016-4215 Sequence 4215, Ap
3	386	18.4	390	3	US-08-905-223-71 Sequence 71, Appl
4	330	15.7	601	4	US-09-949-016-150019 Sequence 150019,
5	330	15.7	35916	4	US-09-949-016-15857 Sequence 15857, A
6	279	13.3	601	4	US-09-949-016-150020 Sequence 150020,
7	189	9.0	601	4	US-09-949-016-150037 Sequence 150037,
8	158	7.5	2475	4	US-09-566-921-88 Sequence 88, Appl
9	155	7.4	601	4	US-09-949-016-150030 Sequence 150030,
10	145	6.9	601	4	US-09-949-016-150031 Sequence 150031,
11	137	6.5	601	4	US-09-949-016-150046 Sequence 150046,
12	137	6.5	601	4	US-09-949-016-150047 Sequence 150047,
13	125	6.0	601	4	US-09-949-016-150029 Sequence 150029,
14	121	5.8	601	4	US-09-949-016-150041 Sequence 150041,
15	121	5.8	601	4	US-09-949-016-150042 Sequence 150042,
16	119	5.7	601	4	US-09-949-016-150008 Sequence 150008,
17	119	5.7	601	4	US-09-949-016-150055 Sequence 150055,
18	110	5.2	601	4	US-09-949-016-150048 Sequence 150048,
19	96	4.5	601	4	US-09-949-016-150032 Sequence 150032,
20	74	3.6	601	4	US-09-949-016-150018 Sequence 150018,
21	69	3.3	244	4	US-09-471-276-495 Sequence 495, App
22	65	3.1	601	4	US-09-949-016-150007 Sequence 150007,
23	30	1.4	1681	4	US-09-023-655-453 Sequence 453, Appl
24	20	1.0	273	4	US-09-513-995C-14761 Sequence 14761, A
25	20	1.0	440	3	US-09-397-787-305 Sequence 305, Appl
26	20	1.0	444	4	US-09-621-976-14139 Sequence 14139, A
27	20	1.0	445	3	US-09-397-787-274 Sequence 274, App

C	28	20	1.0	174259	4	US-09-949-016-11968	Sequence 11968, A
C	29	20	1.0	174262	4	US-09-949-016-14259	Sequence 14259, A
C	30	19	0.9	159	1	US-08-166-346A-8	Sequence 8, Appl
C	31	19	0.9	459	4	US-09-621-976-8324	Sequence 8324, Appl
C	32	19	0.9	3969	3	US-09-518-386B-4	Sequence 4, Appl1
C	33	19	0.9	4396	3	US-09-821-736-1	Sequence 1, Appl1
C	34	19	0.9	14721	4	US-09-949-016-13507	Sequence 13507, A
C	35	19	0.9	25199	4	US-09-949-016-13361	Sequence 13361, A
C	36	19	0.9	129658	4	US-09-949-016-17195	Sequence 17195, A
C	37	19	0.9	186734	4	US-09-949-016-14870	Sequence 14870, A
C	38	19	0.9	193689	4	US-09-949-016-12350	Sequence 12350, A
C	39	19	0.9	193689	4	US-09-949-016-13088	Sequence 13088, A
C	40	19	0.9	200663	4	US-09-949-016-12569	Sequence 12569, A
C	41	19	0.9	203093	4	US-09-949-016-14445	Sequence 14445, A
C	42	18	0.9	78	2	US-08-749-852-56	Sequence 56, Appl
C	43	18	0.9	78	2	US-08-749-852-58	Sequence 58, Appl
C	44	18	0.9	531	4	US-09-252-991A-2223	Sequence 223, Appl
C	45	18	0.9	601	4	US-09-949-016-49781	Sequence 49781, Appl

#### ALIGNMENTS

RESULT 1									
US-09-318-448-23									
; Sequence 23, Application US/09318448									
; Patent No. 6210950									
; GENERAL INFORMATION:									
; APPLICANT: Johnson, William G.									
; APPLICANT: Stenroos, Edward S.									
; TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING									
; TITLE OF INVENTION: DEVELOPMENTAL DISORDERS									
; FILE REFERENCE: 601-1-057									
; CURRENT APPLICATION NUMBER: US/09/318,448									
; CURRENT FILING DATE: 1999-05-25									
; NUMBER OF SEQ ID NOS: 46									
; SOFTWARE: Patent Ver. 2.0									
; SEQ ID NO 23									
; LENGTH: 3259									
; TYPE: DNA									
; ORGANISM: Homo sapiens									
US-09-318-448-23									
Query Match									
Best Local Similarity 97.6%; Score 2046; DB 3; Length 3259;									
Matches 2096; Conservative 0; Mismatches 1; Indels 0; Gaps 0;									
QY	1	ATGAGAGGTTTCTGTTACTATATGCTACACAGCAGGACAGCAAGGCGCATCGAGAA	60						
DB	80	ATGAGAGGTTTCTGTTACTATATGCTACACAGCAGGACAGCAAGGCGCATCGAGAA	139						
QY	61	GAATGTGTGACAGAGCTGTGTACATGATTTTTCGAGATCTTACATATATTAGAA	120						
DB	140	GAATGTGTGACAGAGCTGTGTACATGATTTTTCGAGATCTTACATATATTAGAA	199						
QY	121	TCGGATATGATGACCTAAACCGAAGAGCTCTTGTTGTTGTTTTCACAG	180						
DB	200	TCGGATATGATGACCTAAACCGAAGAGCTCTTGTTGTTGTTTTCACAG	259						
QY	181	GGCAGCGAGACCCAGCCAGCCGCAAGTTGTTAAGAAATACAGCAACA	240						
DB	260	GGCAGCGAGACCCAGCCAGCCGCAAGTTGTTAAGAAATACAGCAACA	319						
QY	241	CTGCGGTTGATTTCTTGTCTACCTGCGGTATGGTTACTGGCTTCGTTATTCAGAA	300						
DB	320	CTGCGGTTGATTTCTTGTCTACCTGCGGTATGGTTACTGGCTTCGTTATTCAGAA	379						
QY	301	TTCACCTACTTTGCAATGGGGGAATATGATTAACGACTTCAAGAGCTTGAGCC	360						
DB	380	TTCACCTACTTTGCAATGGGGGAATATGATTAACGACTTCAAGAGCTTGAGCC	439						
QY	361	CGGCACTTATGACACTGACATGACATGACATGATGATGATTTAGAACTTGTTGAG	420						



Dh 440 CGGATTTCTATGACATGACATGATGATGCTGTAGCTTTAGAACCTTGTGTGAG 439  
Qy 421 CCGTGAATGCTGAGACTCTGCGCAGCCCTGAGAAAGCATTTTAGTTCAGACAGAGCAA 480  
Dh 500 CCGTGATTTCTGGAAGCTCTGGCCAGCCCTGAGAAAGCATTTTAGTTCAGACAGAGCAA 559  
Qy 481 GAGGAGTAAGTGGCGCATCTCCGGTGGCATCACCTGACATCTTGAAGACAGACCTTGTG 540  
Dh 560 GAGGAGTAAGTGGCGCATCTCCGGTGGCATCACCTGACATCTTGAAGACAGACCTTGTG 619  
Qy 541 AAGTCAGAGCTGACATTTGAATCTCAAGTGAAGCTCTGAGATTTCAGATTTAGAGA 600  
Dh 620 AAGTCAGAGCTGACATTTGAATCTCAAGTGAAGCTCTGAGATTTCAGATTTAGAGA 679  
Qy 601 AGAAGAGATCTGAGGTTTGAAGCAAAATGACGTGAACAGCAACCAATCCAAATGTGTGA 660  
Dh 680 AGAAGAGATCTGAGGTTTGAAGCAAAATGACGTGAACAGCAACCAATCCAAATGTGTGA 739  
Qy 661 ATTGAAGATTTGAGTCTTCACTTAACCCGTTCCGATACCCCACTCTCAAGCCTCTCTG 720  
Dh 740 ATTGAAGATTTGAGTCTTCACTTAACCCGTTCCGATACCCCACTCTCAAGCCTCTCTG 799  
Qy 721 AATATTCCTGGTTAAACCCCAAGAAATATTTAAGGTACATCTGACAGAGTCTCTTGGCCAG 780  
Dh 800 AATATTCCTGGTTAAACCCCAAGAAATATTTAAGGTACATCTGACAGAGTCTCTTGGCCAG 859  
Qy 781 GAGGAAAGCCAGATATCTGTGACTTTCAGACAGATTCAGATTTTTCAGAGTCCAAATTTCAAAG 840  
Dh 860 GAGGAAAGCCAGATATCTGTGACTTTCAGACAGATTCAGATTTTTCAGAGTCCAAATTTCAAAG 919  
Qy 841 GCAATTTCACTTACTAGAAATGATGCAATAAAACCACTGCTGGTGAATTTGACATTT 900  
Dh 920 GCAATTTCACTTACTAGAAATGATGCAATAAAACCACTGCTGGTGAATTTGACATTT 979  
Qy 901 TCAATTAAGACCTTTTCTATGAGCCTGAGATGAGCCTTCAAGCCTGATCTGCCCTTAACAGT 960  
Dh 980 TCAATTAAGACCTTTTCTATGAGCCTGAGATGAGCCTTCAAGCCTGATCTGCCCTTAACAGT 1039  
Qy 961 GATTTCTGAGGTACAAAGCCTTACTCCAAAGAATGACGCTTGAAGTAAAGAGAGACCTGC 1020  
Dh 1040 GATTTCTGAGGTACAAAGCCTTACTCCAAAGAATGACGCTTGAAGTAAAGAGAGACCTGC 1099  
Qy 1021 GTCTTTTGAATTAAGGCGACACACAAAGAGAGAGCTTACCCCAAGCATATA 1080  
Dh 1100 GTCTTTTGAATTAAGGCGACACACAAAGAGAGAGCTTACCCCAAGCATATA 1159  
Qy 1081 CCTGCGGAGATGTTCTCCAGTTCATTTTAAACGCTGCTGTAAGTCCGAGCAATTCCT 1140  
Dh 1160 CCTGCGGAGATGTTCTCCAGTTCATTTTAAACGCTGCTGTAAGTCCGAGCAATTCCT 1219  
Qy 1141 AAAAAGGCAATTTTTCGAGGCCCTTGTGACATATACAGTGAACAGTGTGAAAAAGCGCAGG 1200  
Dh 1220 AAAAAGGCAATTTTTCGAGGCCCTTGTGACATATACAGTGAACAGTGTGAAAAAGCGCAGG 1279  
Qy 1201 CTACAGAGAGCTGTGACAGTAAACAAGGGGCGACGCAATTAATGCGGCTTTGATCAAGATGCC 1260  
Dh 1280 CTACAGAGAGCTGTGACAGTAAACAAGGGGCGACGCAATTAATGCGGCTTTGATCAAGATGCC 1339  
Qy 1261 TGTGCGCTGCTGTGATGATCTCCCTGCTGCTTCCCTTGGCAAGCAACCACTCAAGTCTC 1320  
Dh 1340 TGTGCGCTGCTGTGATGATCTCCCTGCTGCTTCCCTTGGCAAGCAACCACTCAAGTCTC 1399  
Qy 1321 CTGCTCGAACAATCTTCTTAACCTTCAACCCAGACCATATTTCTGTGCAAGCTCAAGTTTA 1380  
Dh 1400 CTGCTCGAACAATCTTCTTAACCTTCAACCCAGACCATATTTCTGTGCAAGCTCAAGTTTA 1459  
Qy 1381 TTTCACCCAGAGAAAGCTTCATTTTGTCTTCAACATTTGGAATTTCTGTCTACTGCCACA 1440  
Dh 1460 TTTCACCCAGAGAAAGCTTCATTTTGTCTTCAACATTTGGAATTTCTGTCTACTGCCACA 1519  
Qy 1441 ACAGAGGTTCTGCGAGAGAGATATGACAGGCTGCGCTGCTGTTGTTGCTTCAAGTT 1500  
Dh 1520 ACAGAGGTTCTGCGAGAGAGATATGACAGGCTGCGCTGCTGTTGTTGCTTCAAGTT 1579

Qy 1501 CTTCAAGCAAAACATACATGCATCCCATGAAAGACGCGGAAAAAGCCTGCTCTTAAGATA 1560  
Dh 1580 CTTCAAGCAAAACATACATGCATCCCATGAAAGACGCGGAAAAAGCCTGCTCTTAAGATA 1639  
Qy 1561 TCGATCTCTCTCTGAAACAACAAATTTCTTCCACTTACAGATGACACCCCTCAATCCCCATC 1620  
Dh 1640 TCGATCTCTCTCTGAAACAACAAATTTCTTCCACTTACAGATGACACCCCTCAATCCCCATC 1699  
Qy 1621 ATATAGTGGGATCAGGAAACCGGACATAGCCCGGTTAATGAGGTTCTTCAACAATAGAGAG 1680  
Dh 1700 ATATAGTGGGATCAGGAAACCGGACATAGCCCGGTTAATGAGGTTCTTCAACAATAGAGAG 1759  
Qy 1681 AAATCTCAAGAAACAAACCCAGATGAGAAATTTTGAAGCAATGTGTTTGTGCTGC 1740  
Dh 1760 AAATCTCAAGAAACAAACCCAGATGAGAAATTTTGAAGCAATGTGTTTGTGCTGC 1819  
Qy 1741 AGGCAATAGGATAGGATTTATCTATTCAGAAAAAGACTCAGACATTTCTTAAAGCATGGG 1800  
Dh 1820 AGGCAATAGGATAGGATTTATCTATTCAGAAAAAGACTCAGACATTTCTTAAAGCATGGG 1879  
Qy 1801 ATCTTAATCTACTAAGGTTTCTCTCAAGAGATGCTCTGTGGGAGAGAGAGAGCC 1860  
Dh 1880 ATCTTAATCTACTAAGGTTTCTCTCAAGAGATGCTCTGTGGGAGAGAGAGAGCC 1939  
Qy 1861 CCAGCAAGATATGTACAGAACAAATCCAGCTTCATGACGAGGTTGGCAGAAATCTTC 1920  
Dh 1940 CCAGCAAGATATGTACAGAACAAATCCAGCTTCATGACGAGGTTGGCAGAAATCTTC 1999  
Qy 1921 CTCACAGAGAACGCGCATTTTATGTGTGTGAGATGCAAAAGAAATGTGCCAAGATGTA 1980  
Dh 2000 CTCACAGAGAACGCGCATTTTATGTGTGTGAGATGCAAAAGAAATGTGCCAAGATGTA 2059  
Qy 1981 CATGATGCCCTTGTGCAAAATTAATPAAGAAAGGTTGAGTTGAAAACTAGAAAGCAATG 2040  
Dh 2060 CATGATGCCCTTGTGCAAAATTAATPAAGAAAGGTTGAGTTGAAAACTAGAAAGCAATG 2119  
Qy 2041 AAAACCTGCGCCACTTTAAAGAGAAAAAGCTACCTTCAGAGATTTTGGTCATTA 2097  
Dh 2120 AAAACCTGCGCCACTTTAAAGAGAAAAAGCTACCTTCAGAGATTTTGGTCATTA 2176

RESULT 2  
US-09-949-016-4215  
; Sequence 4215, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: C1001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4215  
; LENGTH: 3242  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-4215

Query Match 90.3%; Score 1893; DB 4; Length 3242;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2093; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGAGAGGTTTCTGTTACTATATGCTACACAGACAGGACAGCAAGGCGATCGCAGAA 60  
|||||

80 ATGAGAGAGTTCTGTATCTATATGCTACACAGCAGGAGCAGCAAAAGCCATCGCAGA 139  
QY 61 GAAATGTGTGACCAAGCTGTGTACATGATTTTCTGCGATCTTCACTATATATGAA 120  
Db 140 GAAATATGTGACCAAGCTGTGTACATGATTTTCTGCGATCTTCACTATATATGAA 199  
QY 121 TCCGATTAAGTATGACCTTAAACCCGAAAGAGCTCTCTGTGTGTGTGTCTACACG 180  
Db 200 TCCGATTAAGTATGACCTTAAACCCGAAAGAGCTCTCTGTGTGTGTGTCTACACG 259  
QY 181 GGCACCGAGAGACCCGACCAACAGAGCCGCAAGTTTGTAAAGAAATACAGAACCAACA 240  
Db 260 GGCACCGAGAGACCCGACCAACAGAGCCGCAAGTTTGTAAAGAAATACAGAACCAACA 319  
QY 241 CTGCGGTGATTTCTTTGCTCACTGCGGTATGAGTTACTGGGTCTCGGTATTCAGAA 300  
Db 320 CTGCGGTGATTTCTTTGCTCACTGCGGTATGAGTTACTGGGTCTCGGTATTCAGAA 379  
QY 301 TACACCTACTTTTGAATGGGGGAGATTAATTGATTAACGACTTCAAGAGCTTGAGCC 360  
Db 380 TACACCTACTTTTGAATGGGGGAGATTAATTGATTAACGACTTCAAGAGCTTGAGCC 439  
QY 361 CGGCAATTTCTATGACACTGAGACATGAGATGACTGTGATGTTTGAACCTTGTTGAG 420  
Db 440 CGGCAATTTCTATGACACTGAGACATGAGATGACTGTGATGTTTGAACCTTGTTGAG 499  
QY 421 CCGTGGATGCTGAGACTGTGCGCAGCCCTCAGAAAGCATTTTATGTCAGCAGAGACA 480  
Db 500 CCGTGGATGCTGAGACTGTGCGCAGCCCTCAGAAAGCATTTTATGTCAGCAGAGACA 559  
QY 481 GAGAGATTAAGTGGGAGCACTCCCGGTGGATCACTGCACTCTTGAGGAGCAGACTTGG 540  
Db 560 GAGAGATTAAGTGGGAGCACTCCCGGTGGATCACTGCACTCTTGAGGAGCAGACTTGG 619  
QY 541 AAGTCAAGCTGTGACATGATGATCTCAAGTCAAGCTTCTGAGATTCAGATTCAGGA 600  
Db 620 AAGTCAAGCTGTGACATGATGATCTCAAGTCAAGCTTCTGAGATTCAGATTCAGGA 679  
QY 601 AGAAAGATTTCTGAGTTTGAACCAAAATGCAATGACAGCAACCAATTCATGTTGTA 660  
Db 680 AGAAAGATTTCTGAGTTTGAACCAAAATGCAATGACAGCAACCAATTCATGTTGTA 739  
QY 661 ATTGAAGACTTGAAGTCTCACTTACCCGTGGTACCCCACTCTCAAGCTCTCTG 720  
Db 740 ATTGAAGACTTGAAGTCTCTACCTTACCCGTGGTACCCCACTCTCAAGCTCTCTG 799  
QY 721 AATAATCTGTGTTTACCCCAAGATATTTTACAGTACATCTGAGAGATCTCTTGCCAG 780  
Db 800 AATAATCTGTGTTTACCCCAAGATATTTTACAGTACATCTGAGAGATCTCTTGCCAG 859  
QY 781 GAGGAAAGCCAAATCTGTGACTTCAAGCATCCAGTTTCAAGTCCCAATTTCAAG 840  
Db 860 GAGGAAAGCCAAATCTGTGACTTCAAGCATCCAGTTTCAAGTCCCAATTTCAAG 919  
QY 841 GCGATTCATCTTCAAGATGATGCTTAAACCACTCTGTGTGTGAATTTGACAT 900  
Db 920 GCGATTCATCTTCAAGATGATGCTTAAACCACTCTGTGTGTGAATTTGACAT 979  
QY 901 TCAATATCAGACTTTCTTCAAGCTGAGATGCTTCAAGCTGATCTGCTTCAAGT 960  
Db 980 TCAATATCAGACTTTCTTCAAGCTGAGATGCTTCAAGCTGATCTGCTTCAAGT 1039  
QY 961 GATTCTGAGTCAAAAGCTTACTCAAAAGCTGAGCTTGAAGATTAAGAGAGCACTGC 1020  
Db 1040 GATTCTGAGTCAAAAGCTTACTCAAAAGCTGAGCTTGAAGATTAAGAGAGCACTGC 1099  
QY 1021 GTCTTTTGAATTAAGGAGACAAAGAAAGAGAGCTTCACTTACCCGACATATA 1080  
Db 1100 GTCTTTTGAATTAAGGAGACAAAGAAAGAGAGCTTCACTTACCCGACATATA 1159  
QY 1081 CTTGCGGAGATGTTCTCTCAGATTCATTTTATCTGTGTCTTGAATTCAGAGCAATTCCT 1140  
Db 1160 CTTGCGGAGATGTTCTCTCAGATTCATTTTATCTGTGTCTTGAATTCAGAGCAATTCCT 1219

QY 1141 AAAAGGCAATTTTTCGAGCCCTTGTGACCTATACAGTACAGTGTCTGAAAAAGCGCAG 1200  
Db 1220 AAAAGGCAATTTTTCGAGCCCTTGTGACCTATACAGTACAGTGTCTGAAAAAGCGCAG 1279  
QY 1201 CTACAGAGCTGTGACATTAACAGAGGCGAGCCGATTAATAGCCGCTTTGACAGATGCC 1260  
Db 1280 CTACAGAGCTGTGACATTAACAGAGGCGAGCCGATTAATAGCCGCTTTGACAGATGCC 1339  
QY 1261 TGTGCTGT 1320  
Db 1340 TGTGCTGT 1399  
QY 1321 CTGCTGAAATCTTCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 1380  
Db 1400 CTGCTGAAATCTTCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 1459  
QY 1381 TTTCAACCGAGAAAGCTCAATTTGTCTTCAATTTGTGAAATTTCTGTCTACCTGACA 1440  
Db 1460 TTTCAACCGAGAAAGCTCAATTTGTCTTCAATTTGTGAAATTTCTGTCTACCTGACA 1519  
QY 1441 ACGAGGTTCTGCGAGAGGAGATGTAAGCTGTGCTGTGTGTGTGTGTGTGTGTGTGT 1500  
Db 1520 ACGAGGTTCTGCGAGAGGAGATGTAAGCTGTGCTGTGTGTGTGTGTGTGTGTGTGT 1579  
QY 1501 CTTCAGCCCAATATCATGCAATCCCATGAAAGCAGGCGGAAAGCCCTGCTCTTAAGATA 1560  
Db 1580 CTTCAGCCCAATATCATGCAATCCCATGAAAGCAGGCGGAAAGCCCTGCTCTTAAGATA 1639  
QY 1561 TCCATCTCTCTCTGACACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
Db 1640 TCCATCTCTCTCTGACACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1699  
QY 1621 ATTAATGT 1680  
Db 1700 ATTAATGT 1759  
QY 1681 AAACTCCAAAGCAACACCCAGATGGAATTTTGAAGCAATGTGTGTGTGTGTGTGTGT 1740  
Db 1760 AAACTCCAAAGCAACACCCAGATGGAATTTTGAAGCAATGTGTGTGTGTGTGTGTGT 1819  
QY 1741 AGGCAATTAAGGATTAAGGATTAATCTTCAAGAAAGAGCTCAGACATTTCTTAAGCATGG 1800  
Db 1820 AGGCAATTAAGGATTAAGGATTAATCTTCAAGAAAGAGCTCAGACATTTCTTAAGCATGG 1879  
QY 1801 ATCTTAATCTTAAGGATTTCTTCTCAAGAGATGCTCTGTGTGTGTGTGTGTGTGTGT 1860  
Db 1880 ATCTTAATCTTAAGGATTTCTTCTCAAGAGATGCTCTGTGTGTGTGTGTGTGTGTGT 1939  
QY 1861 CCAAGCAATTAATTAACAAACATCCAGCTTCAATGCGCAGAGTGGCGAGATCTCTC 1920  
Db 1940 CCAAGCAATTAATTAACAAACATCCAGCTTCAATGCGCAGAGTGGCGAGATCTCTC 1999  
QY 1921 CTTCAGAGAAAGGCGCAATTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1980  
Db 2000 CTTCAGAGAAAGGCGCAATTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2059  
QY 1981 CATGATGCCCTTGTGCAATTAATTAAGCAAAAGGTTGAGTTGAAAACTAGAGCAATG 2040  
Db 2060 CATGATGCCCTTGTGCAATTAATTAAGCAAAAGGTTGAGTTGAAAACTAGAGCAATG 2119  
QY 2041 AAAACCTGTGCCATTTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 2097  
Db 2120 AAAACCTGTGCCATTTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 2176

## RESULT 3

US-08-905-223-71  
; Sequence 71, Application US/08905223  
; Patent No. 6222029  
; GENERAL INFORMATION:  
; APPLICANT: Edwards, Jean-Baptiste D.  
; APPLICANT: Duclert, Aymeric

```

APPLICANT: Lacroix, Bruno
TITLE OF INVENTION: 5' ESTs FOR SECRETED PROTEINS
NUMBER OF SEQUENCES: 503
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 501 West Broadway
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-3505
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Win95
SOFTWARE: Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/905,223
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Israel, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 71:
SEQUENCE CHARACTERISTICS:
LENGTH: 390 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: DOUBLE
TOPOLOGY: LINEAR
MOLECULE TYPE: CDNA
ORIGINAL SOURCE:
ORGANISM: Homo Sapiens
TISSUE TYPE: Brain
FEATURE:
NAME/KEY: s19_peptide
LOCATION: 289..357
IDENTIFICATION METHOD: Von Heijne matrix
OTHER INFORMATION: score 6.9
OTHER INFORMATION: seq SL5LASHSVSC/SN
US-08-905-223-71

Query Match 18.4%; Score 386; DB 3; Length 390;
Best Local Similarity 100.0%; Pred. No. 3.1e-188;
Matches 386; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 970 GTCAAAAGCCTACTCCCAAGACGCTGGAAGATTAAGAGAGCACTGCGCTTTTG 1029
DB 3 GTCAAAAGCCTACTCCCAAGACGCTGGAAGATTAAGAGAGCACTGCGCTTTTG 62
QY 1030 AAATAAAGGACAGACAAAGAAAGAGAGCTACCTTACCCGACATATACCTGGGGA 1089
DB 63 AAATAAAGGACAGACAAAGAAAGAGAGCTACCTTACCCGACATATACCTGGGGA 122
QY 1090 TGTTCCTCAGATTTTAACTGCTGTGTAATTCGAGCAATTCCTAAAAAGCA 1149
DB 123 TGTTCCTCAGATTTTAACTGCTGTGTAATTCGAGCAATTCCTAAAAAGCA 182
QY 1150 TTTTGGAGAGCCCTTGAGACTATACAGTGAAGAGCTGCTGAGAGCTTACAGAG 1209
DB 183 TTTTGGAGAGCCCTTGAGACTATACAGTGAAGAGCTGCTGAGAGCTTACAGAG 242
QY 1210 CTGTCAGTAAACAAGGGGAGCGGATTTAGCCGCTTTGAGAGATGCTGAGCTGC 1269
DB 243 CTGTCAGTAAACAAGGGGAGCGGATTTAGCCGCTTTGAGAGATGCTGAGCTGC 302
QY 1270 TTGTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTGAA 1329
DB 303 TTGTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTGAA 362
QY 1330 CATCTTCCTAAACTCAACCCAGACC 1355

```

```

DB 363 CATCTTCCTAAACTCAACCCAGACC 368
|||||
RESULT 4
US-09-949-016-150019
Sequence 150019, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 150019
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-150019

Query Match 15.7%; Score 330; DB 4; Length 601;
Best Local Similarity 99.7%; Pred. No. 2.2e-159;
Matches 380; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 401 GTTTAGAACTTGTGTTGAGCCGTGATGCTGAGCTTGCGCCAGCCCTCAGAAACATT 460
DB 178 GTTTAGAACTTGTGTTGAGCCGTGATGCTGAGCTTGCGCCAGCCCTCAGAAACATT 237
QY 461 TTAGGTCAGACAGAGACAAAGAGATTAATGCGCACTCCGGTGGCATCCTGCAT 520
DB 238 TTAGGTCAGACAGAGACAAAGAGATTAATGCGCACTCCGGTGGCATCCTGCAT 297
QY 521 CCTTGAGACAGACCTTGGAAGTCAAGCTGCTACATTAATCAAGTGAAGCTTC 580
DB 298 CCTTGAGACAGACCTTGGAAGTCAAGCTGCTACATTAATCAAGTGAAGCTTC 357
QY 581 TGAATTCATGATTCAGAGAAAGAGATTCGAGCTTTGAAGCAAAATGACAGTACA 640
DB 358 TGAATTCATGATTCAGAGAAAGAGATTCGAGCTTTGAAGCAAAATGACAGTACA 417
QY 641 GCAACCAATCCAAATGTTGTAATTAAGACTTTGAGTCTCACTTACCCGTTGGTACCC 700
DB 418 GCAACCAATCCAAATGTTGTAATTAAGACTTTGAGTCTCACTTACCCGTTGGTACCC 477
QY 701 CACTTCAGAGCTCTGGAATATTCCTGTTACCCCGAATATTTACAGGTACATC 760
DB 478 CACTTCAGAGCTCTGGAATATTCCTGTTACCCCGAATATTTACAGGTACATC 537
QY 761 TGCAGAGTCTCTGGCCAGG 781
DB 538 TGCAGAGTCTCTGGCCAGG 558

RESULT 5
US-09-949-016-15957
Sequence 15957, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14

```

PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 15957  
LENGTH: 35916  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-15957

Query Match 15.7%; Score 330; DB 4; Length 35916;  
Best Local Similarity 99.7%; Pred. No. 2.8e-159;  
Matches 380; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 401 GTTTAGAACTTGTGTTGAGCCGTGATGCTGGAAGCTTGGCCAGCCCTCAGAAAGCATT 460  
DB 10781 GTTTAGAACTTGTGTTGAGCCGTGATGCTGGAAGCTTGGCCAGCCCTCAGAAAGCATT 10840  
QY 461 TTAGGTCAAGCAGAGACAAGAGAGATTAAGTGGCCGCACTCCGGTGGCATCACTGCAT 520  
DB 10841 TTAGGTCAAGCAGAGACAAGAGAGATTAAGTGGCCGCACTCCGGTGGCATCACTGCAT 10900  
QY 521 CCTTGAAGCAGACCTTGTGAAGTCAAGCTGCTACACATTAATCTCAAGTGAAGCTTC 580  
DB 10901 CCTTGAAGCAGACCTTGTGAAGTCAAGCTGCTACACATTAATCTCAAGTGAAGCTTC 10960  
QY 581 TGAGATTGATGATTTGAGAAAGAGATTTCTGAGTTTGAAGCAAAATGCAAGTAAACA 640  
DB 10961 TGAGATTGATGATTTGAGAAAGAGATTTCTGAGTTTGAAGCAAAATGCAAGTAAACA 11020  
QY 641 GCAACCAATCCAAATGTTTATTAATTAAGACTTTGAGTCTTCACTTACCCTTGGTACCCC 700  
DB 11021 GCAACCAATCCAAATGTTTATTAATTAAGACTTTGAGTCTTCACTTACCCTTGGTACCCC 11080  
QY 701 CACTCTCACAAGCTCTGTAATATTCCTGTTTACCCCAAGAAATTTTACAGTATCATC 760  
DB 11081 CACTCTCACAAGCTCTGTAATATTCCTGTTTACCCCAAGAAATTTTACAGTATCATC 11140  
QY 761 TGCAGAGTCTCTTGGCCAGG 781  
DB 11141 TGCAGAGTCTCTTGGCCAGG 11161

## RESULT 6

US-09-949-016-150020  
Sequence 150020, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 150020  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-150020

Query Match 13.3%; Score 279; DB 4; Length 601;

Best Local Similarity 99.5%; Pred. No. 4.2e-133;  
Matches 379; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 401 GTTTAGAACTTGTGTTGAGCCGTGATGCTGGAAGCTTGGCCAGCCCTCAGAAAGCATT 460  
DB 165 GTTTAGAACTTGTGTTGAGCCGTGATGCTGGAAGCTTGGCCAGCCCTCAGAAAGCATT 224  
QY 461 TTAGGTCAAGCAGAGACAAGAGAGATTAAGTGGCCGCACTCCGGTGGCATCACTGCAT 520  
DB 225 TTAGGTCAAGCAGAGACAAGAGAGATTAAGTGGCCGCACTCCGGTGGCATCACTGCAT 284  
QY 521 CCTTGAAGCAGACCTTGTGAAGTCAAGCTGCTACACATTAATCTCAAGTGAAGCTTC 580  
DB 285 CCTTGAAGCAGACCTTGTGAAGTCAAGCTGCTACACATTAATCTCAAGTGAAGCTTC 344  
QY 581 TGAGATTGATGATTTGAGAAAGAGATTTCTGAGTTTGAAGCAAAATGCAAGTAAACA 640  
DB 345 TGAGATTGATGATTTGAGAAAGAGATTTCTGAGTTTGAAGCAAAATGCAAGTAAACA 404  
QY 641 GCAACCAATCCAAATGTTTATTAATTAAGACTTTGAGTCTTCACTTACCCTTGGTACCCC 700  
DB 405 GCAACCAATCCAAATGTTTATTAATTAAGACTTTGAGTCTTCACTTACCCTTGGTACCCC 464  
QY 701 CACTCTCACAAGCTCTGTAATATTCCTGTTTACCCCAAGAAATTTTACAGTATCATC 760  
DB 465 CACTCTCACAAGCTCTGTAATATTCCTGTTTACCCCAAGAAATTTTACAGTATCATC 524  
QY 761 TGCAGAGTCTCTTGGCCAGG 781  
DB 525 TGCAGAGTCTCTTGGCCAGG 545

## RESULT 7

US-09-949-016-150037  
Sequence 150037, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 150037  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-150037

Query Match 9.0%; Score 189; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 9.8e-87;  
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1369 AGCTCAAGTTTATTTTACCAGGAAGTCCATTTGTCTTCAACATTTGTAATTTCTG 1428  
DB 18 AGCTCAAGTTTATTTTACCAGGAAGTCCATTTGTCTTCAACATTTGTAATTTCTG 77  
QY 1429 TCTACTGCCACAACAGAGTTCGCGAAGGAGATATGTAAGGCTGCGCTTGTG 1488  
DB 78 TCTACTGCCACAACAGAGTTCGCGAAGGAGATATGTAAGGCTGCGCTTGTG 137  
QY 1489 GTTGCTTCAAGTCTTGAAGCAACATATGATCCATGGAAGACGCGGAAAGCCCTG 1548  
DB 138 GTTGCTTCAAGTCTTGAAGCAACATATGATCCATGGAAGACGCGGAAAGCCCTG 197

Oy	1549	GCTCCTAAG	1557
Db	198	GCTCCTAAG	206

RESULT 8  
US-09-566-921-88  
; Sequence 88, Application US/09566921

	Query Match	7.5%;	Score 158;	DB 4;	Length 2475;
	Best Local Similarity	100.0%;	Pred. No. 9.9e-71;		
	Matches 158;	Conservative	0;	Mismatches 0;	Indels 0;
				Gaps	
QY	525	GAGGACAGACCTTGTGAAGTCAGAGCGTGTAACATTTGAATCTCAAGTCGAGCTTTCGAG			584
DB	16	GAGGACAGACCTTGTGAAGTCAGAGCGTGTAACATTTGAATCTCAAGTCGAGCTTTCGAG			75
QY	585	ATTTCGATGATTCAGGAAGGAAGAGATTCGAGCTTTTGAACCAAAATGCATGGAACAGCA			644
DB	76	ATTTCGATGATTCAGGAAGGAAGAGATTCGAGCTTTTGAACCAAAATGCATGGAACAGCA			135

```

RESULT 9
US-09-949-016-150030
/ Sequence 150030. Application US/09949016
/ Patent No. 681239
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL1001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 150030
/ LENGTH: 601
/ TYPE: DNA
/ ORGANISM: Human
/ US-09-949-016-150030

```

Query Match	7.4%;	Score 155;	DB 4;	Length 601;
Best Local Similarity	100.0%;	Pred. No. 3.2e-69;		
Matches 155;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	904	AATACACACTTTTCCATACAGCCTGGAGATCCCTTCAGCCGATACCTGCCCTTAACAGTAT	963
Db	320	AATACACACTTTTCCATACAGCCTGGAGATCCCTTCAGCCGATACCTGCCCTTAACAGTAT	379
QY	964	TCGAGGTACAAACCCCTACTCCAAAGCTGAGCTTAAAGTAAAGAGACACCTGGCTC	1022
Db	380	TCGAGGTACAAACCCCTACTCCAAAGCTGAGCTTAAAGTAAAGAGACACCTGGCTC	439
QY	1024	CTTTTGAAATTAAGGCGACGACCAAGAACAAAG	1058
Db	440	CTTTTGAAATTAAGGCGACGACCAAGAACAAAG	474

```

RESULT 10
US-09-949-016-150031
: Sequence 150031. Application US/09949016
: Patent No. 6812339
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: C0001307
: CURRENT APPLICATION NUMBER: US/09/949,016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 150031
: LENGTH: 601
: TYPE: DNA
: ORGANISM: Human
US-09-949-016-150031

```

Query Match	6.9%	Score 145;	DB 4;	Length 601;
Best Local Similarity	100.0%;	Pred. No. 4.6e-64;		
Matches 145; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	904	AATACAGCTTTTCCATACAGCTCGAGATACCTTACGCGTGATCGCCCTAAACAGTAT	963
Db	156	AATACAGCTTTTCCATACAGCTCGAGATCGCTTACGCGTGATCGCCCTAAACAGTAT	215
Qy	964	TCGTAGGTACAAACCTTATCCAAAGCTGAGCTTAAAGATAAAGAGACATCGGCTC	1022
Db	216	TCGTAGGTACAAACCTTATCCAAAGCTGAGCTTAAAGATAAAGAGACATCGGCTC	275
Qy	1024	CTTTTGAATAAAGCAGACACAA	1048
Db	276	CTTTTGAATAAAGCAGACACAA	300

```

RESULT 11
US-09-949-016-150046
; Sequence 150046, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL01307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
;

```

```

?
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FASTSEQ for Windows Version 4.0.
? SEQ ID NO 150046
? LENGTH: 601
? TYPE: DNA
? ORGANISM: Human
US-09-949-016-150046

```

Query Match	6.5%	Score 137;	DB 4;	Length 601;	-
Best Local Similarity	99.5%;	Pred. No. 6e-60;			
Matches 187;	Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

Qy	1765	TTGAGAAAAGAGCTCAGACATTTCTTTAAGATGGATCTTAACTCACTTAAAGTTTCC	1824
Db	413	TTGAGAAAAGAGCTCAGACATTTCTTTAAGATGGATCTTAACTCACTTAAAGTTTCC	472
Qy	1825	TTCTCAAGAGATGCTCTCTGTTGGGAGGAGGAAGCCCCAGCAAGATATGTATCAAGACAAC	1884
Db	473	TTCTCAAGAGATGCTCTCTGTTGGGAGGAGGAAGCCCCAGCAAGATATGTATGTCAAGACAAC	532
Qy	1885	ATCCAGCTTCATGGCCAGACAGATGGCCAGAAATCCTCTCCAGAGAAAGGCCATATTTAT	1944
Db	533	ATCCAGCTTCATGGCCAGACAGATGGCCAGAAATCCTCTCCAGAGAAAGGCCATATTTAT	592
Qy	1945	GTGTGTGG 1952	
Db	593	GTGTGTGG 600	

```

RESULT 12
US-09-949-016-150047
; Sequence 150047, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 150047
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-150047

```

	Query Match	Similarity	6.5%	Score 137	DB 4	length 601
	Best Local	Similarity	99.5%	Pred. No.	6e-60	
	Matches	187	Conservative	0	Mismatches	1
					Indels	0
					Gaps	0
Qy	1765	TTCAGAAAAGAGCTCAGACATTTCTTAAGCATGGATCTTAACATCTAAAGGTTCC				1824
Db	191	TTGAGAAAAGAGCTCAGACATTTCTTAAGCATGGATCTTAACATCTAAAGGTTCC				250
Qy	1825	TTCTCAGAGATGCTCTGTGGGGAGAGGAGCCGACGAAGTATGTACAGACAC				1884
Db	251	TTCTCAGAGATGCTCTGTGGGGAGAGGAGCCGACGAAGTATGTACAGACAC				310
Qy	1885	ATTCAGCTTCATGGGCAGCAGGTGGCGAAGTCCCTCCGAGGAGATGGGCATTTAT				1944
Db	311	ATTCAGCTTCATGGGCAGCAGGTGGCGAAGTCCCTCCGAGGAGATGGGCATTTAT				370

QY	1945	GTGTGTGG	1952
Db	371	GTGTGTGG	378

```

RESULT 13
US-09-949-016-150029
; Sequence 150029, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYOMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO010107
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/023,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 150029
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-150029

```

	Query Match	Similarity	6.0%; Score 125; DB 4; Length 601;	
	Best Local	Similarity	100.0%; Pred. No. 9,2e-54;	
	Matches	125; Conservative	0; Mismatches	0; Indels
			0; Gaps	0;
QY	779	AGGAGAAAGCCAGTATCTGTGA	CTTCAGCAGATCCAGTTTTCACGATGGCCAAATTTC	838
Db	379	AGGAGAAAGCCAGTATCTGTGA	CTTCAGCAGATCCAGTTTTCACGATGGCCAAATTTC	438
QY	839	AGGCAGTTCAACTTA	CTACGAATGATGCCATPAAAAACA	CTCTGCTGCGTAGAATTGGACA
Db	439	AGGCAGTTCAACTTA	CTACGAATGATGCCATPAAAAACA	CTCTGCTGCGTAGAATTGGACA
QY	899	TTTGCA	903	
Db	499	TTTGCA	503	

```

RESULT 14
US-09-949-016-150041
; Sequence 150041: Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 150041
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-150041

```

Query Match	5.88;	Score 121;	DB 4;	Length 601;
-------------	-------	------------	-------	-------------

Best Local Similarity 100.0%; Pred. No. 1.1e-51;  
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1556 AGATATCCATCTCTCTCGAACAACAATTCCTTTCACCTTACGATGACCCCTCAATCC 1615

Db 124 AGATATCCATCTCTCTCGAACAACAATTCCTTTCACCTTACGATGACCCCTCAATCC 183

QY 1616 CCATCATTAATGGTGGGTCCAGGAACCGGCATAGCCCGTTTATTGGGTTCTTACAACATA 1675

Db 184 CCATCATTAATGGTGGGTCCAGGAACCGGCATAGCCCGTTTATTGGGTTCTTACAACATA 243

QY 1676 G 1676

Db 244 G 244

## RESULT 15

US-09-949-016-150042

; Sequence 150042, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: C1001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 150042

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-150042

Query Match 5.8%; Score 121; DB 4; Length 601;

Best Local Similarity 100.0%; Pred. No. 1.1e-51;

Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1556 AGATATCCATCTCTCTCGAACAACAATTCCTTTCACCTTACGATGACCCCTCAATCC 1615

Db 95 AGATATCCATCTCTCTCGAACAACAATTCCTTTCACCTTACGATGACCCCTCAATCC 154

QY 1616 CCATCATTAATGGTGGGTCCAGGAACCGGCATAGCCCGTTTATTGGGTTCTTACAACATA 1675

Db 155 CCATCATTAATGGTGGGTCCAGGAACCGGCATAGCCCGTTTATTGGGTTCTTACAACATA 214

QY 1676 G 1676

Db 215 G 215

Search completed: August 27, 2005, 16:18:18  
Job time : 237.757 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: August 27, 2005, 00:17:56 ; Search time 900.401 Seconds  
(without alignments)  
15238.608 Million cell updates/sec

Title: US-09-371-347a-43

Sequence: 1 atgagagaggttcctgtaact.....ttcagagatattggtcaca 2097

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 7331713 seqs, 3271544945 residues

Word size : 0

Total number of hits satisfying chosen parameters: 14663426

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Published Applications NA.\*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US05\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq.\*  
20: /cgn2\_6/ptodata/2/pubpna/US10H\_PUBCOMB.seq.\*  
21: /cgn2\_6/ptodata/2/pubpna/US10I\_PUBCOMB.seq.\*  
22: /cgn2\_6/ptodata/2/pubpna/US11A\_PUBCOMB.seq.\*  
23: /cgn2\_6/ptodata/2/pubpna/US11B\_PUBCOMB.seq.\*  
24: /cgn2\_6/ptodata/2/pubpna/US11C\_PUBCOMB.seq.\*  
25: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
26: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2097	100.0	2097	10	US-09-371-347-43
2	2046	97.6	2097	10	US-09-371-347-1
3	2046	97.6	3259	10	US-09-371-347-24
4	1995	95.1	2097	10	US-09-371-347-41
5	1803	86.0	2094	10	US-09-371-347-45
6	1742	83.1	2093	10	US-09-371-347-47
7	1018	48.5	3256	21	US-10-741-600-692

8	1018	48.5	3274	21	US-10-741-600-693	Sequence 693, App
9	330	15.7	591	16	US-10-029-386-6369	Sequence 6369, Ap
10	328	15.6	379	16	US-10-029-386-20100	Sequence 20100, A
11	279	13.3	591	16	US-10-029-386-1735	Sequence 1735, Ap
12	277	13.2	379	16	US-10-029-386-15435	Sequence 15435, A
13	266	12.7	43985	21	US-10-741-600-17757	Sequence 17757, A
14	188	9.0	525	16	US-10-029-386-653	Sequence 653, App
15	175	8.3	175	16	US-10-029-386-14338	Sequence 14338, A
16	158	7.5	2475	10	US-09-909-5678-38	Sequence 38, App1
17	158	7.5	2475	22	US-10-765-760-88	Sequence 88, App1
18	158	7.5	21852	21	US-10-741-600-17986	Sequence 17986, A
19	150	7.2	201	21	US-10-741-600-15583	Sequence 15583, A
20	150	7.2	201	21	US-10-741-600-15584	Sequence 15584, A
21	150	7.2	201	21	US-10-741-600-15589	Sequence 15589, A
22	150	7.2	201	21	US-10-741-600-15590	Sequence 15590, A
23	150	7.2	201	21	US-10-741-600-15592	Sequence 15592, A
24	150	7.2	201	21	US-10-741-600-15593	Sequence 15593, A
25	150	7.2	201	21	US-10-741-600-15594	Sequence 15594, A
26	150	7.2	201	21	US-10-741-600-15598	Sequence 15598, A
27	150	7.2	201	21	US-10-741-600-15599	Sequence 15599, A
28	150	7.2	201	21	US-10-741-600-15600	Sequence 15600, A
29	150	7.2	201	21	US-10-741-600-15602	Sequence 15602, A
30	150	7.2	201	21	US-10-741-600-15606	Sequence 15606, A
31	150	7.2	201	21	US-10-741-600-15609	Sequence 15609, A
32	150	7.2	201	21	US-10-741-600-15610	Sequence 15610, A
33	150	7.2	201	21	US-10-741-600-15612	Sequence 15612, A
34	150	7.2	201	21	US-10-741-600-15613	Sequence 15613, A
35	150	7.2	201	21	US-10-741-600-15614	Sequence 15614, A
36	150	7.2	201	21	US-10-741-600-15620	Sequence 15620, A
37	150	7.2	201	21	US-10-741-600-15621	Sequence 15621, A
38	150	7.2	201	21	US-10-741-600-15623	Sequence 15623, A
39	150	7.2	201	21	US-10-741-600-15624	Sequence 15624, A
40	150	7.2	201	21	US-10-741-600-15625	Sequence 15625, A
41	150	7.2	201	21	US-10-741-600-15629	Sequence 15629, A
42	150	7.2	201	21	US-10-741-600-15630	Sequence 15630, A
43	150	7.2	201	21	US-10-741-600-15631	Sequence 15631, A
44	150	7.2	201	21	US-10-741-600-15633	Sequence 15633, A
45	150	7.2	201	21	US-10-741-600-15637	Sequence 15637, A

#### ALIGNMENTS

RESULT 1  
US-09-371-347-43  
; Sequence 43, Application US/09371347  
; Publication No. US20030082676A1  
; GENERAL INFORMATION:  
; APPLICANT: Roy A. Gravel et al.  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE REDUCTASE;  
; TITLE OF INVENTION: CLONING, AND METHODS FOR EVALUATING RISK OF NEURAL TUBE  
; FILE REFERENCE: 50004/003003  
; CURRENT FILING DATE: 1999-08-10  
; PRIOR FILING DATE: 1999-08-10  
; PRIOR APPLICATION NUMBER: 60/071,622  
; PRIOR FILING DATE: 1998-01-16  
; PRIOR APPLICATION NUMBER: 09/232,028  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 43  
; LENGTH: 2097  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-371-347-43

Query Match 100.0%; Score 2097; DB 10; Length 2097;  
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;  
Matches 2097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 ATGAGAGGTTCTGTTACTATATGCTACACAGCAGGACAGGCAAGGCATCGCAGAA 60  
|||||



TITLE OF INVENTION: CLONING, AND METHODS FOR EVALUATING RISK OF NEURAL TUBE  
FILE REFERENCE: 50004/001003  
CURRENT APPLICATION NUMBER: US/09/371,347  
CURRENT FILING DATE: 1999-08-10  
PRIOR APPLICATION NUMBER: 60/071,622  
PRIOR FILING DATE: 1998-01-16  
PRIOR APPLICATION NUMBER: 09/232,028  
PRIOR FILING DATE: 1999-01-15  
NUMBER OF SEQ ID NOS: 51  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 2097  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-371-347-1

Query Match 97.6%; Score 2046; DB 10; Length 2097;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2096; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGAGAGTTTCTGTACTATATGCTACACAGGAGGACAGGCAAGGCCATCGCAGAA 60  
DB 1 ATGAGAGAGTTTCTGTACTATATGCTACACAGGAGGACAGGCAAGGCCATCGCAGAA 60  
QY 61 GAAATGTGAGCAAGCTGTGTACATGATTTTCTGAGATCTTCACTATATTAGTGA 120  
DB 61 GAAATGTGAGCAAGCTGTGTACATGATTTTCTGAGATCTTCACTATATTAGTGA 120  
QY 121 TCCGATAGTATGACCTTAACAAACGAAAGCTCTCTGTGTGTGTGTCTACACG 180  
DB 121 TCCGATAGTATGACCTTAACAAACGAAAGCTCTCTGTGTGTGTGTCTACACG 180  
QY 181 GGCACCGGAGACCCACCCGACACAGCCCGCAAGTTTGTATGAAATACAAACCA 240  
DB 181 GGCACCGGAGACCCACCCGACACAGCCCGCAAGTTTGTATGAAATACAAACCA 240  
QY 241 CTGCGGTTGATTTCTTGTCTACCTGCGGTATGAGTTCTGGGTCGCGTATTCAGAA 300  
DB 241 CTGCGGTTGATTTCTTGTCTACCTGCGGTATGAGTTCTGGGTCGCGTATTCAGAA 300  
QY 301 TACACCTACTTTTGCATGAGGGGGAAGATTAATTGATTAACGACTTCAGAGCTTGAGCC 360  
DB 301 TACACCTACTTTTGCATGAGGGGGAAGATTAATTGATTAACGACTTCAGAGCTTGAGCC 360  
QY 361 CGGCAATTTCTATGACACTGACATGCAATGACTGTGATGATTAGAACTTGTGTTGAG 420  
DB 361 CGGCAATTTCTATGACACTGACATGCAATGACTGTGATGATTAGAACTTGTGTTGAG 420  
QY 421 CCGTGAATGCTGGAATCTGCGCAGCCCTCAGAAACATTTTATGTCAGAGAGAGCAA 480  
DB 421 CCGTGAATGCTGGAATCTGCGCAGCCCTCAGAAACATTTTATGTCAGAGAGAGCAA 480  
QY 481 GAGGAGATAGTGGCCCACTCCCGGTGCACTGCACTCTTGAGAGCAGACCTTGTG 540  
DB 481 GAGGAGATAGTGGCCCACTCCCGGTGCACTGCACTCTTGAGAGCAGACCTTGTG 540  
QY 541 AAGTCAGAGCTGCTACATTTGAATCAAGTCAAGCTTCTGAGATTCGATGATTCAGGA 600  
DB 541 AAGTCAGAGCTGCTACATTTGAATCAAGTCAAGCTTCTGAGATTCGATGATTCAGGA 600  
QY 601 AGAAAGATTTCTGAGGTTTGAAGCAAAATGCAAGTAAACAGCAACCAATTCATGTTGA 660  
DB 601 AGAAAGATTTCTGAGGTTTGAAGCAAAATGCAAGTAAACAGCAACCAATTCATGTTGA 660  
QY 661 ATGAAGACTTTGAGTCTCACTTACCCGTTCCGTTACCCCACTCTCAAGGCTCTCTG 720  
DB 661 ATGAAGACTTTGAGTCTCACTTACCCGTTCCGTTACCCCACTCTCAAGGCTCTCTG 720  
QY 721 AATATTCCTGTTTACCCCAAGATATTTACAGGTACATCTGACAGAGTCTCTTGGCCAG 780  
DB 721 AATATTCCTGTTTACCCCAAGATATTTACAGGTACATCTGACAGAGTCTCTTGGCCAG 780

QY 781 GAGAAAGCCAGATATCTGTGACTTCAGAGATTCAGATTTTCAAGTCCAAATTTCAAG 840  
DB 781 GAGAAAGCCAGATATCTGTGACTTCAGAGATTCAGATTTTCAAGTCCAAATTTCAAG 840  
QY 841 GCAATTCATCTTACGAGATATGCAATTAACCACTCTGCTGTATGATTTGACATT 900  
DB 841 GCAATTCATCTTACGAGATATGCAATTAACCACTCTGCTGTATGATTTGACATT 900  
QY 901 TCAATATCACTTTTCTATCAGGCTGAGATGCTTCAAGGTGATCTGCCCTAACAT 960  
DB 901 TCAATATCACTTTTCTATCAGGCTGAGATGCTTCAAGGTGATCTGCCCTAACAT 960  
QY 961 GATTCGAGTACAAAGCTTCTCAAGAGCTGACAGTTGAAGATTAAGAGAGCATG 1020  
DB 961 GATTCGAGTACAAAGCTTCTCAAGAGCTGACAGTTGAAGATTAAGAGAGCATG 1020  
QY 1021 GTCTTTGAAAATTAAGGACACACAAAGAAAGAGCTTACCTTACCCAGCATTA 1080  
DB 1021 GTCTTTGAAAATTAAGGACACACAAAGAAAGAGCTTACCTTACCCAGCATTA 1080  
QY 1081 CCGGCGGAGATGTTCTCTCAGTTCAATTTTACCTGCTGTTGAATCCGAGCAATTC 1140  
DB 1081 CCGGCGGAGATGTTCTCTCAGTTCAATTTTACCTGCTGTTGAATCCGAGCAATTC 1140  
QY 1141 AAAAAGCATTTTTCGAGCCCTTGTGACATATACAGTACAGTGTGAAAAAGCCAG 1200  
DB 1141 AAAAAGCATTTTTCGAGCCCTTGTGACATATACAGTACAGTGTGAAAAAGCCAG 1200  
QY 1201 CTACAGAGCTGTGACATTAACAGGGGAGCCGATTAAGCCGCTTTGACAGATGCC 1260  
DB 1201 CTACAGAGCTGTGACATTAACAGGGGAGCCGATTAAGCCGCTTTGACAGATGCC 1260  
QY 1261 TGTGCTGTGTGATGATCTCTCTGCTGCTTCCCTTCTGCAAGCACAACCTCACTC 1320  
DB 1261 TGTGCTGTGTGATGATCTCTCTGCTGCTTCCCTTCTGCAAGCACAACCTCACTC 1320  
QY 1321 CTGCTGAAATCTTCTTAACTTCAACCCAGACATATTCGTGACAGTCAAGTTTA 1380  
DB 1321 CTGCTGAAATCTTCTTAACTTCAACCCAGACATATTCGTGACAGTCAAGTTTA 1380  
QY 1381 TTTCAACCGAAGAGTCAATTTTGTCTTCAATTTGAAATTTCTGTACTGACACA 1440  
DB 1381 TTTCAACCGAAGAGTCAATTTTGTCTTCAATTTGAAATTTCTGTACTGACACA 1440  
QY 1441 ACAGAGTTCTGCGAAGGAGATATGACAGCTGCGCTTGTGATGCTTCAAGTT 1500  
DB 1441 ACAGAGTTCTGCGAAGGAGATATGACAGCTGCGCTTGTGATGCTTCAAGTT 1500  
QY 1501 CTTGAGCCAAATATCATGATCCCATGAAAGACAGGGGAAAGCCCTGCTCTTAAGATA 1560  
DB 1501 CTTGAGCCAAATATCATGATCCCATGAAAGACAGGGGAAAGCCCTGCTCTTAAGATA 1560  
QY 1561 TCCATCTCTCTGAAACAAATTTTCTCACTTACAGATGACCCCTCAATCCCATC 1620  
DB 1561 TCCATCTCTCTGAAACAAATTTTCTCACTTACAGATGACCCCTCAATCCCATC 1620  
QY 1621 ATATATGTTGTTCTGAGAACCGGCAATGCGCTTATTTGAGTTCTTAAACATTAAG 1680  
DB 1621 ATATATGTTGTTCTGAGAACCGGCAATGCGCTTATTTGAGTTCTTAAACATTAAG 1680  
QY 1681 AAATCTCAAGAACCAACCCAGATGGAATTTTGAAGCATGTGTTTGTGCTGC 1740  
DB 1681 AAATCTCAAGAACCAACCCAGATGGAATTTTGAAGCATGTGTTTGTGCTGC 1740  
QY 1741 AGCATTAAGATTAAGATTAATCTATTCAGAAAAGCTCAGACATTTCTTAAAGATG 1800  
DB 1741 AGCATTAAGATTAAGATTAATCTATTCAGAAAAGCTCAGACATTTCTTAAAGATG 1800  
QY 1801 ATCTTAATCACTTAAGATTTCTTCTCAAGAGATGCTCTGTTGGGAGAGAGAGCC 1860  
DB 1801 ATCTTAATCACTTAAGATTTCTTCTCAAGAGATGCTCTGTTGGGAGAGAGAGCC 1860  
QY 1861 CCAGCAAGATATGTAACAAGAACATTCAGTTTCAATGCGACAGAGTGGCAGAAATCTC 1920

Db 1861 CGAGCAAGATATATACAAAGCAATCCAGCTTCATGCGCAGCAGGGGAGAAATCTTC 1920  
Qy 1921 CTCAGAGAAAGGACCATATATTTATGTGTGAGATCAAAAGATATGCGCAAGATGTA 1980  
Db 1921 CTCAGAGAAAGGACCATATATTTATGTGTGAGATCAAAAGATATGCGCAAGATGTA 1980  
Qy 1981 CATGATGCCCTTGTGCAAAATATATAGCAAAAGAGTTGAGTTGAAAACTAGAACATG 2040  
Db 1981 CATGATGCCCTTGTGCAAAATATATAGCAAAAGAGTTGAGTTGAAAACTAGAACATG 2040  
Qy 2041 AAAACCTGGCCCTTTAAAAAGAAAGAAACGCTACCTTCAGATATTTGGTATTA 2097  
Db 2041 AAAACCTGGCCCTTTAAAAAGAAAGAAACGCTACCTTCAGATATTTGGTATTA 2097  
  
RESULT 3  
US-09-371-347-24  
Sequence 24, Application US/09371347  
Publication No. US2003082676A1  
GENERAL INFORMATION:  
APPLICANT: Roy A. Gravel et al.  
TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE REDUCTASE:  
TITLE OF INVENTION: CLONING, AND METHODS FOR EVALUATING RISK OF NEURAL TUBE  
TITLE OF INVENTION: DEFECTS, CARDIOVASCULAR DISEASE, AND CANCER  
FILE REFERENCE: 50004/003003  
CURRENT APPLICATION NUMBER: US/09/371.347  
PRIOR FILING DATE: 1999-08-10  
PRIOR APPLICATION NUMBER: 60/071.622  
PRIOR FILING DATE: 1998-01-16  
PRIOR APPLICATION NUMBER: 09/232.028  
NUMBER OF SEQ ID NOS: 51  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 24  
LENGTH: 3259  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-371-347-24  
  
Query Match 97.6%; Score 2046; DB 10; Length 3259;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2096; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 ATGAGGAGGTTTCTGTACTATATGCTTACACAGCAGGAGCAGCAAGGCCATCGCAGAA 60  
Db 80 ATGAGGAGGTTTCTGTACTATATGCTTACACAGCAGGAGCAGCAAGGCCATCGCAGAA 139  
Qy 61 GAATATGTGAGCAAGCTGTGTGATCATGATTTTCTGCAATCTTCACTATATTAAGTAA 120  
Db 140 GAATATGTGAGCAAGCTGTGTGATCATGATTTTCTGCAATCTTCACTATATTAAGTAA 199  
Qy 121 TCCGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180  
Db 200 TCCGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 259  
Qy 181 GGACCGGAGACCCACCCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAAACAAACA 240  
Db 260 GGACCGGAGACCCACCCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAAACAAACA 319  
Qy 241 CTGCGCGGTTATTTCTTGTCTCACTGCGGTATGAGTTTACTGGGTCTCGGTGATTCAGAA 300  
Db 320 CTGCGCGGTTATTTCTTGTCTCACTGCGGTATGAGTTTACTGGGTCTCGGTGATTCAGAA 379  
Qy 301 TACACCTACTTTTGGCAATGGGGGAGATATTAAGTAAAGCACTTCAAGAGCTTGGAGCC 360  
Db 380 TACACCTACTTTTGGCAATGGGGGAGATATTAAGTAAAGCACTTCAAGAGCTTGGAGCC 439  
Qy 361 CGGCAATTTCTATGACATGACATGACATGACATGACATGACATGACATGACATGACATG 420  
Db 440 CGGCAATTTCTATGACATGACATGACATGACATGACATGACATGACATGACATGACATG 499  
Qy 421 CCGTGATTTGCTGAGCTCTGGCCAGCCCTCAGAAAGCAATTTTAGGTCAACAGACAGCAA 480

Db 500 CCGTGATTTGCTGAGCTCTGGCCAGCCCTCAGAAAGCAATTTTAGGTCAACAGACAGCAA 559  
Qy 481 GAGAGATATAGTGGCCACCTCCGGTGGCATCACTGCACTCTTGAAGACAGACCTTGTG 540  
Db 560 GAGAGATATAGTGGCCACCTCCGGTGGCATCACTGCACTCTTGAAGACAGACCTTGTG 619  
Qy 541 AAGTCAGAGCTGTACATATGATCTCAAGTGCAGCTTCTGAGATTCGATGATTCAGGA 600  
Db 620 AAGTCAGAGCTGTACATATGATCTCAAGTGCAGCTTCTGAGATTCGATGATTCAGGA 679  
Qy 601 AGAAAGATTTCTGAGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATCCATGTTGTA 660  
Db 680 AGAAAGATTTCTGAGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATCCATGTTGTA 739  
Qy 661 ATTAAGACTTTGAGTCTCATCTTACCCGTTGGGTACCCCACTCTCACAGGCTCTGTG 720  
Db 740 ATTAAGACTTTGAGTCTCATCTTACCCGTTGGGTACCCCACTCTCACAGGCTCTGTG 799  
Qy 721 AATATTCCTGTTTACCCCGAATATTTACAGGTACATCTGAGAGAGTCTTGGCCAG 780  
Db 800 AATATTCCTGTTTACCCCGAATATTTACAGGTACATCTGAGAGAGTCTTGGCCAG 859  
Qy 781 GAGGAAAGCCAGATCTGTGATGACATCCAGTTTTCAGATGCGCAATTTGCAAG 840  
Db 860 GAGGAAAGCCAGATCTGTGATGACATCCAGTTTTCAGATGCGCAATTTGCAAG 919  
Qy 841 GCAGTTCAACTTACTACGATGATGCAATTAACCACTCTGCTGTAGATTTGACAT 900  
Db 920 GCAGTTCAACTTACTACGATGATGCAATTAACCACTCTGCTGTAGATTTGACAT 979  
Qy 901 TCAATATCAGACTTTTCTATCAGCTCTGAGATGCTTCAAGCTGATCTGCTTACAGT 960  
Db 980 TCAATATCAGACTTTTCTATCAGCTCTGAGATGCTTCAAGCTGATCTGCTTACAGT 1039  
Qy 961 GATTTGAGGTAAAGCCCTACCCCAAGACCTGAGCTTGAAGTAAAGAGGACCTGC 1020  
Db 1040 GATTTGAGGTAAAGCCCTACCCCAAGACCTGAGCTTGAAGTAAAGAGGACCTGC 1099  
Qy 1021 GTCTTTTGAATTAAGGACAGACACAAAGAAAGAGCTACTTACCCCGACATATA 1080  
Db 1100 GTCTTTTGAATTAAGGACAGACACAAAGAAAGAGCTACTTACCCCGACATATA 1159  
Qy 1081 CTGCGGAGATTTCTCTCAAGTTCATTTTACTGCTGTCTTGAATTCGAGCAATTCCT 1140  
Db 1160 CTGCGGAGATTTCTCTCAAGTTCATTTTACTGCTGTCTTGAATTCGAGCAATTCCT 1219  
Qy 1141 AAAAAGCAATTTTGGAGCCCTTGTGCACTATACAGAGACAGTCTGAAAAGCCGAGG 1200  
Db 1220 AAAAAGCAATTTTGGAGCCCTTGTGCACTATACAGAGACAGTCTGAAAAGCCGAGG 1279  
Qy 1201 CTACAGAGCTGTGCAATTAAGGAGGAGCCGATTAATGAGCGCTTGTGAGAGATGCC 1260  
Db 1280 CTACAGAGCTGTGCAATTAAGGAGGAGCCGATTAATGAGCGCTTGTGAGAGATGCC 1339  
Qy 1261 TGTGCTGCTTGTGATCTCTCTGCTTCTTCTTCTTCCATGAGCCACCACTCACTC 1320  
Db 1340 TGTGCTGCTTGTGATCTCTCTGCTTCTTCTTCTTCCATGAGCCACCACTCACTC 1399  
Qy 1321 CTGCTGGAACATCTTCTAACTTCAACCCAGACATATTCGTGTCAGAGCTCAAGTTTA 1380  
Db 1400 CTGCTGGAACATCTTCTAACTTCAACCCAGACATATTCGTGTCAGAGCTCAAGTTTA 1459  
Qy 1381 TTTCAACCCAGGAAGCTCAATTTTGTCTTCAACATTTGGAATTTGTCTACTGACACA 1440  
Db 1460 TTTCAACCCAGGAAGCTCAATTTTGTCTTCAACATTTGGAATTTGTCTACTGACACA 1519  
Qy 1441 ACAGAGTTCTGCGAAGGAGATATGACAGCTGAGCTGCTTGTGTTGCTTCAATT 1500  
Db 1520 ACAGAGTTCTGCGAAGGAGATATGACAGCTGAGCTGCTTGTGTTGCTTCAATT 1579  
Qy 1501 CTTCAGCCAAACATATATGATCCCATGGAAGACAGGCGGAAAGCCCTGCTCTTAAGATA 1560

Db	1580	CTTCAGGCAAAACATACATTCGATCCCATGAGAAAGAGGCGGAAAGGCCCTGGCTCCTAAGATA	1639
Qy	1561	TCATCTCTCTCTCGAACAACAATTCTTTCCACTTACCAAGATGACCCCTCAATCCCATC	1620
Db	1640	TCCATCTCTCTCTCGAACAACAATTCTTTCCACTTACCAAGATGACCCCTCAATCCCATC	1699
Qy	1621	ATATGATGGTCCAGAACCCGATAGCCCCCTTTATGAGTTCCTACAAATAGAGAG	1680
Db	1700	ATAATGATGGTCCAGAACCCGATAGCCCCCTTTATGAGTTCCTACAAATAGAGAG	1759
Qy	1681	AAACTCCAGAACCAACCCAGATGGAATTTTGGAGCAATGCTGTTGTTTTGGCTGC	1740
Db	1760	AAACTCCAGAACCAACCCAGATGGAATTTTGGAGCAATGCTGTTGTTTTGGCTGC	1819
Qy	1741	AGGCATAGAGTATGGAGTTATCTATTCCAGAAAAGAGCTCAGACATTTCTTTAAGCATGCG	1800
Db	1820	AGGCATAGAGTATGGAGTTATCTATTCCAGAAAAGAGCTCAGACATTTCTTTAAGCATGCG	1879
Qy	1801	ATCTTAACTCATCTAAGAGTTTCTCTTCAGAGATGCTCCTGTTGGGAGAGAAAGCC	1860
Db	1880	ATCTTAACTCATCTAAGAGTTTCTCTTCAGAGATGCTCCTGTTGGGAGAGAAAGCC	1939
Qy	1861	CCAGCAAAAGTATGTACAGAACAAATCCAGTTTATGCGCCAGAGGTGGCGAATCTCTC	1920
Db	1940	CCAGCAAAAGTATGTACAGAACAAATCCAGTTTATGCGCCAGAGGTGGCGAATCTCTC	1999
Qy	1921	CTCCAGAGAAACGGCCATATTATTATGTTGTTGAGATGCAAGAAATATGCGCCAGAGATGTA	1980
Db	2000	CTCCAGAGAAACGGCCATATTATTATGTTGTTGAGATGCAAGAAATATGCGCCAGAGATGTA	2059
Qy	1981	CATCATGCCCTTGTGCAATATATTAAGCAAGAGGTGGAGTTGAAAACTAGAAGCATG	2040
Db	2060	CATCATGCCCTTGTGCAATATATTAAGCAAGAGGTGGAGTTGAAAACTAGAAGCATG	2119
Qy	2041	AAAACCTGCGCACTTTAAAGAAAGAAACGCTACCTTCAGATATTTGGTCATPA	2097
Db	2120	AAAACCTGCGCACTTTAAAGAAAGAAACGCTACCTTCAGATATTTGGTCATPA	2176
RESULT 4			
US-09-371-347-41			
Sequence 41, Application US/09371347			
Publication No. US20030082676A1			
GENERAL INFORMATION:			
APPLICANT: Roy A. Gravel et al.			
TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE REDUCTASE:			
TITLE OF INVENTION: CLONING, AND METHODS FOR EVALUATING RISK OF NEURAL TUBES			
FILE REFERENCE: 50004/003003			
CURRENT APPLICATION NUMBER: US/09/371,347			
CURRENT FILING DATE: 1999-08-10			
PRIOR APPLICATION NUMBER: 60/071,622			
PRIOR FILING DATE: 1998-01-16			
PRIOR APPLICATION NUMBER: 09/232,028			
PRIOR FILING DATE: 1999-01-15			
NUMBER OF SEQ ID NOS: 51			
SOFTWARE: FastSeq for Windows Version 4.0			
SEQ ID NO 41			
LENGTH: 2097			
TYPE: DNA			
ORGANISM: Homo sapiens			
US-09-371-347-41			

Query	Match	Similarity	95.1%	Score	1995;	DB	10;	Length	2097;
	Best Local	Similarity	99.9%	Pred.	No. 0;				
	Matches	2095;	Conservative	0;	Mismatches	2;	Indels	0;	Gaps
									0;
Qy	1	ATGAGGAGGTTCTCTTACTATATGCTTACAGCAGGAGGACAGGAGCCATGCGACAA	60						
Db	1	ATGAGGAGGTTCTCTTACTATATGCTTACAGCAGGAGGACAGGAGCCATGCGACAA	60						
Qy	61	GAAATGTTGTAGCAGAGCTGTGTGATCATGATTTCTGCAGATCTTCACTATATTATGTAA	120						

D	61	GAATATATGTAGCAAGCTGTGTGTATCATGTGATTTTCTGCAAGATCTTCACTGTAATTATGAAA	120
Q	121	TCGGATTAAGTATGACTTAAAAACCGAAACAGCTCCTCTGTGTGTGTGTCTTATCAACG	180
D	121	TCCGATTAAGTATGACTTAAAAACCGAAACAGCTCCTCTGTGTGTGTGTCTTATCAACG	180
Q	181	GGCAACGGAGACCCACCGCACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACA	240
D	181	GGCAACGGAGACCCACCGCACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACA	240
Q	241	CTGCGGTGTGATTTCTTTGTCTCACTTCGCGTATGGGTAACTGGGTCTCGGTGATTCAGAA	300
D	241	CTGCGGTGTGATTTCTTTGTCTCACTTCGCGTATGGGTAACTGGGTCTCGGTGATTCAGAA	300
Q	301	TACACCTTACTTTTGCAATGGGGGGGAAGATTAATTGATTAACGACTTCAAGACTTGAGCC	360
D	301	TACACCTTACTTTTGCAATGGGGGGGAAGATTAATTGATTAACGACTTCAAGACTTGAGCC	360
Q	361	CGGCAATTTCTATGACATCTGACATGCGAGATGACCTGTGTAGGTTTGAACCTTGTGGTAG	420
D	361	CGGCAATTTCTATGACATCTGACATGCGAGATGACCTGTGTAGGTTTGAACCTTGTGGTAG	420
Q	421	CCGTGATTTGTGGAAGCTGTGGCCAGACCCTCAGAAAGACTTTTAGTCAAGCAGACAGCAA	480
D	421	CCGTGATTTGTGGAAGCTGTGGCCAGACCCTCAGAAAGACTTTTAGTCAAGCAGACAGCAA	480
Q	481	GAGAGATTAAGTGGCGCATCTCCGGTGGCATCACCTGCAATCTTTGAGACAGACTTTGTG	540
D	481	GAGAGATTAAGTGGCGCATCTCCGGTGGCATCACCTGCAATCTTTGAGACAGACTTTGTG	540
Q	541	AAGTCAGAGCTGTCAACATTTGAATCTTCAGTCCAGATTCGATGATTCAGGA	600
D	541	AAGTCAGAGCTGTCAACATTTGAATCTTCAGTCCAGATTCGATGATTCAGGA	600
Q	601	AGAAAGGATTTCTGAGTTTGAAGCAAAATGCAGTGAACAGCAACCAATCCAATGTGTGA	660
D	601	AGAAAGGATTTCTGAGTTTGAAGCAAAATGCAGTGAACAGCAACCAATCCAATGTGTGA	660
Q	661	ATTGAGAATTTGAGTCTCTCACTTACCCTGTGGTACCCCACTTCAAGAGCTCTCTG	720
D	661	ATTGAGAATTTGAGTCTCTCACTTACCCTGTGGTACCCCACTTCAAGAGCTCTCTG	720
Q	721	AATATTCCTGTTTACCCCGAATATTTACAGGTACATCTGCAGAGATCTCTTGGCCAG	780
D	721	AATATTCCTGTTTACCCCGAATATTTACAGGTACATCTGCAGAGATCTCTTGGCCAG	780
Q	781	GAGGAAAGCCAAAGTATCTGTGACTTTCAGACATCCAGTTTTCAGAGTCCAAATTTCAAG	840
D	781	GAGGAAAGCCAAAGTATCTGTGACTTTCAGACATCCAGTTTTCAGAGTCCAAATTTCAAG	840
Q	841	GCAGTTCAACTTACTAGCAATGATGCCATAAAACACTCTGCTGTGAAATTGACAATT	900
D	841	GCAGTTCAACTTACTAGCAATGATGCCATAAAACACTCTGCTGTGAAATTGACAATT	900
Q	901	TCAATATCAGACTTTTCTCTATCAGCTGTGAGATGCTTCAAGCTGATCTGCCCTTACAGT	960
D	901	TCAATATCAGACTTTTCTCTATCAGCTGTGAGATGCTTCAAGCTGATCTGCCCTTACAGT	960
Q	961	GATTCGTAGGTATCAAAAGCTTACTCCAAAGATTCGAAAGATTAAGAGAGACATGCG	1020
D	961	GATTCGTAGGTATCAAAAGCTTACTCCAAAGATTCGAAAGATTAAGAGAGACATGCG	1020
Q	1021	GTCTCTTTGAAATTAAGGACAGACAAAGAAAGAGAGCTTACTTACCCACACATATA	1080
D	1021	GTCTCTTTGAAATTAAGGACAGACAAAGAAAGAGAGCTTACTTACCCACACATATA	1080
Q	1081	CTGCGGAGATTTCTCTCAAGTTCAATTTTACCTGTGTCTTGAATTCGAGCAATTTCT	1140
D	1081	CTGCGGAGATTTCTCTCAAGTTCAATTTTACCTGTGTCTTGAATTCGAGCAATTTCT	1140
Q	1141	AAAAAGGATTTTTCGAGCCCTTGTGACTTATCAGTGAACAGTGCAGAAAGGCGCAGG	1200
D	1141	AAAAAGGATTTTTCGAGCCCTTGTGACTTATCAGTGAACAGTGCAGAAAGGCGCAGG	1200

QY 1201 CTACAGAGTGTGCAAGTAAACAGGGGACGCCGATTATAGCCGCTTTGTACGAGATGCC 1260  
DB 1201 CTACAGAGTGTGCAAGTAAACAGGGGACGCCGATTATAGCCGCTTTGTACGAGATGCC 1260  
QY 1261 TGTGCGCTGTGTGATCTCTCTCGCTTTCCCTTTCCCTTTCCGAGCAGACACTGATCTC 1320  
DB 1261 TGTGCGCTGTGTGATCTCTCTCGCTTTCCCTTTCCCTTTCCGAGCAGACACTGATCTC 1320  
QY 1321 CTGCTGGAACATCTTCTTAACCTTCAACCCAGACATATTCGTGTGAGCTCAAGTTTA 1380  
DB 1321 CTGCTGGAACATCTTCTTAACCTTCAACCCAGACATATTCGTGTGAGCTCAAGTTTA 1380  
QY 1381 TTTTACCAGGAAAGCTCCATTTTGTCTTCAACATGTGGAATTTCTGTCTACTGCCACA 1440  
DB 1381 TTTTACCAGGAAAGCTCCATTTTGTCTTCAACATGTGGAATTTCTGTCTACTGCCACA 1440  
QY 1441 ACAGAGTTTCTGGGAGAGGAGATATGACAGGCTGGCTGGCTTTGTTGTTGTTCAAGTT 1500  
DB 1441 ACAGAGTTTCTGGGAGAGGAGATATGACAGGCTGGCTGGCTTTGTTGTTGTTCAAGTT 1500  
QY 1501 CTTCAGCCAAACATACATGATCCCATGAGACAGCGGAAAGCCCTGCTCTCTTAAGATA 1560  
DB 1501 CTTCAGCCAAACATACATGATCCCATGAGACAGCGGAAAGCCCTGCTCTCTTAAGATA 1560  
QY 1561 TCCATCTCTCTGGAACACAAATTTTTCACCTTACAGATGACCCCTCAATCCCATC 1620  
DB 1561 TCCATCTCTCTGGAACACAAATTTTTCACCTTACAGATGACCCCTCAATCCCATC 1620  
QY 1621 ATATAGTGTGGTCCAGGACCGGATGACCCGCTTATTTGGGTTTCCACAACTAAGAGAG 1680  
DB 1621 ATATAGTGTGGTCCAGGACCGGATGACCCGCTTATTTGGGTTTCCACAACTAAGAGAG 1680  
QY 1681 AAATCCAAAGAACACACCCAGATGGAATTTTGGAGCAATGTGTGTTTGGCTGC 1740  
DB 1681 AAATCCAAAGAACACACCCAGATGGAATTTTGGAGCAATGTGTGTTTGGCTGC 1740  
QY 1741 AGGATATAGATAGGATTAATCTATTCAGAAAGAGCTCAGACATTTCTTAAAGCATGG 1800  
DB 1741 AGGATATAGATAGGATTAATCTATTCAGAAAGAGCTCAGACATTTCTTAAAGCATGG 1800  
QY 1801 ATCTTAACTATCTAAAGGTTTCTTCTTCAAGAGATCTCTGTGGGGAGAGAGAGCC 1860  
DB 1801 ATCTTAACTATCTAAAGGTTTCTTCTTCAAGAGATCTCTGTGGGGAGAGAGAGCC 1860  
QY 1861 CCAGCAAGATATGACAAACATCAGCTTCAATGAGCAGAGAGTGGGAGATCTC 1920  
DB 1861 CCAGCAAGATATGACAAACATCAGCTTCAATGAGCAGAGAGTGGGAGATCTC 1920  
QY 1921 CTCAGAGAGAACGCGCATATTTATGTGTGAGAGTCAAGAAATATGCGCAAGATGTA 1980  
DB 1921 CTCAGAGAGAACGCGCATATTTATGTGTGAGAGTCAAGAAATATGCGCAAGATGTA 1980  
QY 1981 CAGATATCTTGTGCAAAATTAAGCAAGAGTGTGAGTGAAGAACTAAGAGATG 2040  
DB 1981 CAGATATCTTGTGCAAAATTAAGCAAGAGTGTGAGTGAAGAACTAAGAGATG 2040  
QY 2041 AAAACCTGGCACTTTAAAGAGAAAGAGCTACCTTCAAGATATTTGTCTATA 2097  
DB 2041 AAAACCTGGCACTTTAAAGAGAAAGAGCTACCTTCAAGATATTTGTCTATA 2097

RESULT 5  
US-09-371-347-45  
; Sequence 45, Application US/09371347  
; Publication No. us2003082676A1  
; GENERAL INFORMATION:  
; APPLICANT: Roy A. Gravel et al.  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE REDUCTASE;  
; TITLE OF INVENTION: CLONING, AND METHODS FOR EVALUATING RISK OF NEURAL TUBE  
; TITLE OF INVENTION: DEFECTS, CARDIOVASCULAR DISEASE, AND CANCER  
; FILE REFERENCE: 50004/003003  
; CURRENT APPLICATION NUMBER: US/09/371,347

; CURRENT FILING DATE: 1999-08-10  
; PRIOR APPLICATION NUMBER: 60/071,622  
; PRIOR FILING DATE: 1998-01-16  
; PRIOR APPLICATION NUMBER: 09/232,028  
; PRIOR FILING DATE: 1999-01-15  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 45  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-371-347-45

Query Match 86.0%; Score 1803; DB 10; Length 2094;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2093; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 1 ATGAGAGAGTTTCTGTACTATATGCTACACAGGAGGACAGGCAAGGCCATCGCAGAA 60  
DB 1 ATGAGAGAGTTTCTGTACTATATGCTACACAGGAGGACAGGCAAGGCCATCGCAGAA 60  
QY 61 GAAATGTGAGCAAGCTGTGTACATGATTTTCTGAGATCTTCAATATTTAGTGA 120  
DB 61 GAAATGTGAGCAAGCTGTGTACATGATTTTCTGAGATCTTCAATATTTAGTGA 120  
QY 121 TCCGATATGATGACCTTAAACCGAAGAGCTCTTGTGTGTGTTTACACAG 180  
DB 121 TCCGATATGATGACCTTAAACCGAAGAGCTCTTGTGTGTGTTTACACAG 180  
QY 181 GGCACGGAGACCCACCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACAA 240  
DB 181 GGCACGGAGACCCACCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACAA 240  
QY 241 CTGCGGTGATTTCTTGTCTCACTGCGGTATGAGTTAAGGTTCTGGTATTCAGA 300  
DB 241 CTGCGGTGATTTCTTGTCTCACTGCGGTATGAGTTAAGGTTCTGGTATTCAGA 300  
QY 301 TACACCTACTTTTGCAATGGGGGAAAGATATGATTAACGACTTCAAGAGCTTGA 360  
DB 301 TACACCTACTTTTGCAATGGGGGAAAGATATGATTAACGACTTCAAGAGCTTGA 360  
QY 361 CGGATTTCTAAGACACTGGAACATGAGATGACTGTGTAGAGTTAGAACTGTGTTG 420  
DB 361 CGGATTTCTAAGACACTGGAACATGAGATGACTGTGTAGAGTTAGAACTGTGTTG 420  
QY 421 CCGTGAATGTGAGACTGTGCGCAGCTCCGATGATCACTGCACTTGTGAGAGACG 480  
DB 421 CCGTGAATGTGAGACTGTGCGCAGCTCCGATGATCACTGCACTTGTGAGAGACG 480  
QY 481 GAGGAGATATGAGGCGCACTCCGATGATCACTGCACTTGTGAGAGACG 540  
DB 481 GAGGAGATATGAGGCGCACTCCGATGATCACTGCACTTGTGAGAGACG 540  
QY 541 AATCGAGCTGTACACATGAAATCTCAAGTGAAGCTTGTGAGTTGATTCAGAA 600  
DB 541 AATCGAGCTGTACACATGAAATCTCAAGTGAAGCTTGTGAGTTGATTCAGAA 600  
QY 601 AGAAGGATTTGAGGTTTGAAGCAAAATGACGTAACAGCAACCAATCCAATGTTGA 660  
DB 601 AGAAGGATTTGAGGTTTGAAGCAAAATGACGTAACAGCAACCAATCCAATGTTGA 660  
QY 661 ATTGAAGCTTTGAGTCTTCACTTACCCGTTCCGTTACCCCACTCTCAAGCTCTCTG 720  
DB 661 ATTGAAGCTTTGAGTCTTCACTTACCCGTTCCGTTACCCCACTCTCTCAAGCTCTCTG 720  
QY 721 AATATTCGTGTTTACCCCGAGAAATTTTACAGGTATCATGAGAGTCTTGGCAG 780  
DB 721 AATATTCGTGTTTACCCCGAGAAATTTTACAGGTATCATGAGAGTCTTGGCAG 780  
QY 781 GAGGAAAGCAAGTATCTGATCTTCAAGATCAAGTATTTCAAGTCCAAATTTCAAG 840  
DB 781 GAGGAAAGCAAGTATCTGATCTTCAAGATCAAGTATTTCAAGTCCAAATTTCAAG 840







Db 481 GAGAGGTAAGTGGCGCATCCCGGTGGCATCTGCACTTGGAGACAGACTTGTG 540  
 Qy 541 AAGTCAGAGCTGCTACATTTGATCTCAAGTCAGAGCTTGGATTCATGATTCAGGA 600  
 Db 541 AAGTCAGAGCTGCTACATTTGATCTCAAGTCAGAGCTTGGATTCATGATTCAGGA 600  
 Qy 601 AGAAGAGATTCTGAGGTTTGAAGCAAAATGCAAGTGAACCAACCAATTCATGTTGA 660  
 Db 601 AGAAGAGATTCTGAGGTTTGAAGCAAAATGCAAGTGAACCAACCAATTCATGTTGA 660  
 Qy 661 ATTGAAGACTTTGAGTCTCACTTACCCGTTGGTACCCCACTTCACAAAGCTCTCG 720  
 Db 661 ATTGAAGACTTTGAGTCTCACTTACCCGTTGGTACCCCACTTCACAAAGCTCTCG 720  
 Qy 721 AATATTCCTGGTTTACCCCAAGATATTTACAGGTATCTGACAGAGTCTTTGGCCAG 780  
 Db 721 AATATTCCTGGTTTACCCCAAGATATTTACAGGTATCTGACAGAGTCTTTGGCCAG 780  
 Qy 781 GAGGAAAGCCAGATATCTGATCTTCAAGATCTCAAGTCTTCAAGTCCAAATTCAAAG 840  
 Db 781 GAGGAAAGCCAGATATCTGATCTTCAAGATCTCAAGTCTTCAAGTCCAAATTCAAAG 840  
 Qy 841 GCAATTCACCTTACATGAGATGATGCAATTAACCACTCTGCTGGTGAATTCGACAT 900  
 Db 841 GCAATTCACCTTACATGAGATGATGCAATTAACCACTCTGCTGGTGAATTCGACAT 900  
 Qy 901 TCAATTCACAGACTTTCTCTACAGCTGAGATGCTTCAAGCTTCAAGCTTCAAGCT 960  
 Db 901 TCAATTCACAGACTTTCTCTACAGCTGAGATGCTTCAAGCTTCAAGCTTCAAGCT 960  
 Qy 961 GATTCTGAGGTACAAAGCTTCTCAAAAGCTGACGCTTGAAGTAAAGAGACATGC 1020  
 Db 961 GATTCTGAGGTACAAAGCTTCTCAAAAGCTGACGCTTGAAGTAAAGAGACATGC 1020  
 Qy 1021 GTCTCTTTGAAATTAAGGACACACAAAGAAAGAGACTTACCTTCCCGCATATA 1080  
 Db 1021 GTCTCTTTGAAATTAAGGACACACAAAGAAAGAGACTTACCTTCCCGCATATA 1080  
 Qy 1081 CTTGCGGAGATGTTCTCTCCAGATCATTTTACCTGGTCTTGAATTCGAGCAATTCCT 1140  
 Db 1081 CTTGCGGAGATGTTCTCTCCAGATCATTTTACCTGGTCTTGAATTCGAGCAATTCCT 1140  
 Qy 1141 AAAAAGGCAATTTTTCGAGCCCTTGTGACATATCAAGTACAGTGTGAAAAGCCGAG 1200  
 Db 1141 AAAAAGGCAATTTTTCGAGCCCTTGTGACATATCAAGTACAGTGTGAAAAGCCGAG 1200  
 Qy 1201 CTACAGAGAGCTGTCAGTAAACAAAGGGGACCGATTAATAGCCGCTTTGACAGATGCC 1260  
 Db 1201 CTACAGAGAGCTGTCAGTAAACAAAGGGGACCGATTAATAGCCGCTTTGACAGATGCC 1260  
 Qy 1261 TGTGCTGCTGTTGATCTCTCTGCTCTGCTTCCCTTCTTGCCAGCCACATCAAGTCTC 1320  
 Db 1261 TGTGCTGCTGTTGATCTCTCTGCTCTGCTTCCCTTCTTGCCAGCCACATCAAGTCTC 1320  
 Qy 1321 CTGCTGCAACATCTTCTTAACTTCAACCCAGACCAATATTCGATGCAAGCTCAAGTTTA 1380  
 Db 1321 CTGCTGCAACATCTTCTTAACTTCAACCCAGACCAATATTCGATGCAAGCTCAAGTTTA 1380  
 Qy 1381 TTTTCACCCAAGAAAGCTTCAATTTTGTCTTCAACATTTGTGAAATTTCTGTCTACGCA 1440  
 Db 1381 TTTTCACCCAAGAAAGCTTCAATTTTGTCTTCAACATTTGTGAAATTTCTGTCTACGCA 1440  
 Qy 1441 ACAGAGGTTCTGCGAAGGAGATGTAACAGGCTGAGCTTGTGTTGCTTCAAGT 1500  
 Db 1441 ACAGAGGTTCTGCGAAGGAGATGTAACAGGCTGAGCTTGTGTTGCTTCAAGT 1500  
 Qy 1501 CTTTACGCAAAACATATACATGATCCCATGAGAGACGCGGAAAGCCCTGCTCTTAAGATA 1560  
 Db 1501 CTTTACGCAAAACATATACATGATCCCATGAGAGACGCGGAAAGCCCTGCTCTTAAGATA 1560  
 Qy 1561 TCCATCTCTCTCTGAAACAAATTTCTTTCATTTACATGATGACCCCTCAATCCCATC 1620  
 Db 1561 TCCATCTCTCTCTGAAACAAATTTCTTTCATTTACATGATGACCCCTCAATCCCATC 1620

Db 1561 TCCATCTCTCTCTGAAACAAATTTCTTTCATTTACATGATGACCCCTCAATCCCATC 1620  
 Qy 1621 ATATATGAGGTCACAGAACCCGACATAGCCCGTTTATTTGGTTCTTACATATAGAG 1680  
 Db 1621 ATATATGAGGTCACAGAACCCGACATAGCCCGTTTATTTGGTTCTTACATATAGAG 1680  
 Qy 1681 AAATCTCAAGAACACACCCAGATGGAATTTTGGACATGTGTTTGTGTTGGCTGC 1740  
 Db 1681 AAATCTCAAGAACACACCCAGATGGAATTTTGGACATGTGTTTGTGTTGGCTGC 1740  
 Qy 1741 AAGCATTAAGATTAAGGATTAATCTATTCAGAAAAGACTCAGACATTTCTTAAAGATGG 1800  
 Db 1741 AAGCATTAAGATTAAGGATTAATCTATTCAGAAAAGACTCAGACATTTCTTAAAGATGG 1800  
 Qy 1797 ATCTTAATCTCACTTAAGGTTTCTTCTCAAGAGATGCTCTGTGGGAGAGAAAGCC 1856  
 Db 1797 ATCTTAATCTCACTTAAGGTTTCTTCTCAAGAGATGCTCTGTGGGAGAGAAAGCC 1856  
 Qy 1861 CCAGCAAGATTAATGTAACAGACATCCAGCTTTCATGCGCAGAGGTGGCAGAAATCTTC 1920  
 Db 1861 CCAGCAAGATTAATGTAACAGACATCCAGCTTTCATGCGCAGAGGTGGCAGAAATCTTC 1920  
 Qy 1921 CTCACAGAGAACGCGCATTTATGATGATGAGATGCAAAAGATATGGCCAAAGATGTA 1980  
 Db 1921 CTCACAGAGAACGCGCATTTATGATGATGAGATGCAAAAGATATGGCCAAAGATGTA 1980  
 Qy 1981 CATGATGCTCTGTCGAAATTAATTAAGCAAGGTTGAGATTGGAATTAAGCAATG 2040  
 Db 1981 CATGATGCTCTGTCGAAATTAATTAAGCAAGGTTGAGATTGGAATTAAGCAATG 2040  
 Qy 2041 AAAACCTTGCCACTTTAAAGAAAGAAACGCTTACCTCAGATATTTGGTATATA 2097  
 Db 2041 AAAACCTTGCCACTTTAAAGAAAGAAACGCTTACCTCAGATATTTGGTATATA 2097  
 Qy 2097 AAAACCTTGCCACTTTAAAGAAAGAAACGCTTACCTCAGATATTTGGTATATA 2097  
 Db 2097 AAAACCTTGCCACTTTAAAGAAAGAAACGCTTACCTCAGATATTTGGTATATA 2097

RESULT 7  
 US-10-741-600-692  
 ; Sequence 692, Application US/10741600  
 ; Publication No. US20050026169A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CARGILL, Michele et al.  
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
 ; FILE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF  
 ; CURRENT APPLICATION NUMBER: US/10/741,600  
 ; CURRENT FILING DATE: 2003-12-22  
 ; NUMBER OF SEQ ID NOS: 73997  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 692  
 ; LENGTH: 3256  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-741-600-692

Query Match 48.5%; Score 1018; DB 21; Length 3256;  
 Best Local Similarity 99.0%; Pred. No. 0;  
 Matches 1968; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 111 TATTAGATATCGATTAAGTATGATGATTAAGAAACGAAACGATCTCTTGTGTTGGT 170  
 Db 204 TATTAGATATCGATTAAGTATGATGATTAAGAAACGAAACGATCTCTTGTGTTGGT 263  
 Qy 171 TTCTACACGCGGACCGAGACCCACCGACACAGCCCGCAAGTTTGTAAAGAAATACA 230  
 Db 264 TTCTACACGCGGACCGAGACCCACCGACACAGCCCGCAAGTTTGTAAAGAAATACA 323  
 Qy 231 GAAACAAACCTGCGGTTGATTTCTTGTCTACCTGCGGATATGGTTACTGGGTCTCG 290  
 Db 324 GAAACAAACCTGCGGTTGATTTCTTGTCTACCTGCGGATATGGTTACTGGGTCTCG 383  
 Qy 291 TGAATGAAATTAACCTACTTTTGAATGGGGGAAAGATTAATGATTAAGCACTTCAAGA 350  
 Db 384 TGAATGAAATTAACCTACTTTTGAATGGGGGAAAGATTAATGATTAAGCACTTCAAGA 443

351 GCTTGGAGCCCGGCACTTTCTATGACA CTGGAGACATGAGATGACTGTTAGTTAGT 410  
444 GCTTGGAGCCCGGCACTTTCTATGACA CTGGAGACATGAGATGACTGTTAGTTAGT 503  
411 TGTGTTGAGCCGTGATGCTGAGACTGCGCAGCCCTCAGAAACATTTTAGTCAAG 470  
504 TGTGTTGAGCCGTGATGCTGAGACTGCGCAGCCCTCAGAAACATTTTAGTCAAG 563  
471 CAGAGGACAAAGAGATTAAGTGGCCACTCCCGTGGCATCACTGCAATCTTGGAGAC 530  
564 CAGAGGACAAAGAGATTAAGTGGCCACTCCCGTGGCATCACTGCAATCTTGGAGAC 623  
531 AGACCTTGGAAGTGAAGCTGCTACACATGAAATCTCAAGTGAAGCTTCTGAGATTCGA 590  
624 AGACCTTGGAAGTGAAGCTGCTACACATGAAATCTCAAGTGAAGCTTCTGAGATTCGA 683  
591 TGAATCAGAGAGAAAGATTTCTGAGGTTTGAAGCAAAATGCAATGACAGCAACCAATC 650  
684 TGAATCAGAGAGAAAGATTTCTGAGGTTTGAAGCAAAATGCAATGACAGCAACCAATC 743  
651 CAATGTTGTAATTGAAGACTTGAAGTCTCACTTACCCGTTGGTACCCCACTCTCACA 710  
744 CAATGTTGTAATTGAAGACTTGAAGTCTCACTTACCCGTTGGTACCCCACTCTCACA 803  
711 AGCCTCTGGAATATCTGCTGTTTACCCCGAATATTTACAGTTCATCTGAGAGATC 770  
804 AGCCTCTGGAATATCTGCTGTTTACCCCGAATATTTACAGTTCATCTGAGAGATC 863  
771 TCTTGGCCAGAGAGAAAGCCAAATATCTGATCTGAGCAGATCCAGTTTTCAGTGGCC 830  
864 TCTTGGCCAGAGAGAAAGCCAAATATCTGATCTGAGCAGATCCAGTTTTCAGTGGCC 923  
831 AATTCAAAAGCAGTTCAACTTACAGATGATGATGATGATGATGATGATGATGATGATGAT 890  
924 AATTCAAAAGCAGTTCAACTTACAGATGATGATGATGATGATGATGATGATGATGATGAT 983  
891 ATTGGAATTTCAAAATACAGATCTTTCATGACCTGAGAGATGCTTCAAGCTGATCTG 950  
984 ATTGGAATTTCAAAATACAGATCTTTCATGACCTGAGAGATGCTTCAAGCTGATCTG 1043  
951 CCTTACAGATGATTTGAGGATCAAAAGCTTCTCAAAAGCTGAGCTGAGATGATGATGAT 1010  
1044 CCTTACAGATGATTTGAGGATCAAAAGCTTCTCAAAAGCTGAGCTGAGATGATGATGAT 1103  
1011 AGAGCACTGGCTCTTTTGAATAAAGGAGACACAAAGAGAAAGAGCTTACC 1070  
1104 AGAGCACTGGCTCTTTTGAATAAAGGAGACACAAAGAGAAAGAGCTTACC 1163  
1071 CCAGCATATACCTGCGGAGATTTCTCCAGTTCAATTTTACCTGCTGTTGAAATCCG 1130  
1164 CCAGCATATACCTGCGGAGATTTCTCCAGTTCAATTTTACCTGCTGTTGAAATCCG 1223  
1131 AGCAATTTCTTAAAGAGCATTTTTCGAGGCTTTCGAGCTTATACAGATGAGTGTGA 1190  
1224 AGCAATTTCTTAAAGAGCATTTTTCGAGGCTTTCGAGCTTATACAGATGAGTGTGA 1283  
1191 AAAGCCAGGCTTACAGAGAGCTGTCAGATTAACAAGGGGAGCCGATTTATAGCCGCTTGT 1250  
1284 AAAGCCAGGCTTACAGAGAGCTGTCAGATTAACAAGGGGAGCCGATTTATAGCCGCTTGT 1343  
1251 AGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1310  
1344 AGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1403  
1311 ACTCAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1370  
1404 ACTCAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1463  
1371 CTCAAGTTTATTTTCAACCAGAGAAAGCTCAATTTTGTCTTCAACATGTTGAAATTTCTGTC 1430  
1464 CTCAAGTTTATTTTCAACCAGAGAAAGCTCAATTTTGTCTTCAACATGTTGAAATTTCTGTC 1523

1431 TACTGCCACAGAGAGTTCTGCGAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1490  
1524 TACTGCCACAGAGAGTTCTGCGAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1583  
1491 TGTCTCAGTTCTTCAAGCCAAACATATGATGATGATGATGATGATGATGATGATGATGATGAT 1550  
1584 TGTCTCAGTTCTTCAAGCCAAACATATGATGATGATGATGATGATGATGATGATGATGATGAT 1643  
1551 TGTCTCAGTTCTTCAAGCCAAACATATGATGATGATGATGATGATGATGATGATGATGATGAT 1610  
1644 TGTCTCAGTTCTTCAAGCCAAACATATGATGATGATGATGATGATGATGATGATGATGATGAT 1703  
1611 AATCCCATCATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1670  
1704 AATCCCATCATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1763  
1671 ACATAGAGAGAACTTCAAGAAACAACCCAGATGAAATTTTGGAGCAATGTTGTT 1730  
1764 ACATAGAGAGAACTTCAAGAAACAACCCAGATGAAATTTTGGAGCAATGTTGTT 1823  
1731 TTTTGGCTGAGCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1790  
1824 TTTTGGCTGAGCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1883  
1791 TTAAGCATGAGATCTTATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1850  
1884 TTAAGCATGAGATCTTATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1943  
1851 GGAGAGAGCCCGCAGCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1910  
1944 GGAGAGAGCCCGCAGCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2003  
1911 GAGATCTCTCTCCAG 1970  
2004 GAGATCTCTCTCCAG 2063  
1971 CAAGAT 2030  
2064 CAAGAT 2123  
2031 AGAAGCATGAAACCTTGCCCACTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2090  
2124 AGAAGCATGAAACCTTGCCCACTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2183  
2091 GTCATTA 2097  
2184 GTCATTA 2190

RESULT 8  
US-10-741-600-693  
; Sequence 693, Application US/10741600  
; Publication No. US20050026169A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001499  
; CURRENT APPLICATION NUMBER: US/10/741,600  
; CURRENT FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 73997  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 693  
; LENGTH: 3274  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-741-600-693

Query Match 48.5%; Score 1018; DB 21; Length 3274;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 1968; Conservative 0; Mismatches 19; Indels 0; Gaps 0;  
111 TATTAGTGAATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 170

|||||  
Db 222 TATTAGTAATCCGATAGTATGACTTAAACCAGAAACAGCTCTCTTGTGTGGT 281  
Qy 171 TTCTACACCGGGACCCGAGACCCACGACAGCCCGCAAGTTGTTAAGAAATACA 230  
Db 282 TTCTACACGGGGACCGGAGACCCACGACAGCCCGCAAGTTGTTAAGAAATACA 341  
Qy 231 GAACCAAACTCGCCGGTGAATTTCTTGTCTCACCTGCGGTATGGTTACTGGCTCGG 230  
Db 342 GAACCAAACTCGCCGGTGAATTTCTTGTCTCACCTGCGGTATGGTTACTGGCTCGG 401  
Qy 291 TGAATTCAGATACACTCTTTTGCATATGGGGGAGATAATTAATGATTAACGACTTCAAGA 350  
Db 402 TGATTCAGATACACTCTTTTGCATATGGGGGAGATAATTAATGATTAACGACTTCAAGA 461  
Qy 351 GCTTGAAGCCCGCATTTCTATGACACTGACATGACATGACTGTGTAGTTAGAACT 410  
Db 462 GCTTGAAGCCCGCATTTCTATGACACTGACATGACATGACTGTGTAGTTAGAACT 521  
Qy 411 TGTGTGAGCCCGTGAATTTGCTGACCTGTGCGCCAGCCCTCAAGAAAGATTTTAGTCAAG 470  
Db 522 TGTGTGAGCCCGTGAATTTGCTGACCTGTGCGCCAGCCCTCAAGAAAGATTTTAGTCAAG 581  
Qy 471 CAGAGCAAGAGAGATTAAGTGGCGGACCTCCGCTGGCATCACTGATCTCTTGAAGAC 530  
Db 582 CAGAGCAAGAGAGATTAAGTGGCGGACCTCCGCTGGCATCACTGATCTCTTGAAGAC 641  
Qy 531 AGACTTGTGATGACAGAGCTGTACACATTTGAATCTCAAGTGAAGCTTGTGAGATTCGA 590  
Db 642 AGACTTGTGATGACAGAGCTGTACACATTTGAATCTCAAGTGAAGCTTGTGAGATTCGA 701  
Qy 591 TGATTCAGAGAAAGAGATTTCTGAGTTTGAAGCAAAATGCAAGTGAACGCAACCAATC 650  
Db 702 TGATTCAGAGAAAGAGATTTCTGAGTTTGAAGCAAAATGCAAGTGAACGCAACCAATC 761  
Qy 651 CAATGTGTATTAAGAAGCTTTGAGTCTCACTTACCCGTTCCGGTACCCCACTGTACA 710  
Db 762 CAATGTGTATTAAGAAGCTTTGAGTCTCACTTACCCGTTCCGGTACCCCACTGTACA 821  
Qy 711 AGCTCTCTGATATATCTCTGCTTTACCCCAAGATATTTACAGGTACATCTGACAGAGTC 770  
Db 822 AGCTCTCTGATATATCTCTGCTTTACCCCAAGATATTTACAGGTACATCTGACAGAGTC 881  
Qy 771 TCTTGGCAGAGAGAAAGCCAGATATCTGACTTCAAGCATCTCAAGTTTCAAGTGGC 830  
Db 882 TCTTGGCAGAGAGAAAGCCAGATATCTGACTTCAAGCATCTCAAGTTTCAAGTGGC 941  
Qy 831 AATTCAGAGAGAGTTCAACTTACTACGAATGATGCCATTAACCACTGCTGCTGATGA 890  
Db 942 AATTCAGAGAGAGTTCAACTTACTACGAATGATGCCATTAACCACTGCTGCTGATGA 1001  
Qy 891 ATTGGAATTTCAAAATACAGCTTTTCTATCAGCTGAGATGCTTCAAGCTGATCTG 950  
Db 1002 ATTGGAATTTCAAAATACAGCTTTTCTATCAGCTGAGATGCTTCAAGCTGATCTG 1061  
Qy 951 CCCCTAAGATGATTTCTGAGTACAAAGCTTACTCAAAAGATGAGCTTGAAGTAAAG 1010  
Db 1062 CCCCTAAGATGATTTCTGAGTACAAAGCTTACTCAAAAGATGAGCTTGAAGTAAAG 1121  
Qy 1011 AGAGCACTGCGCTCTTTTGAATAATTAAGCAGACACAAAGAAAGAGACTACCTTACC 1070  
Db 1122 AGAGCACTGCGCTCTTTTGAATAATTAAGCAGACACAAAGAAAGAGACTACCTTACC 1181  
Qy 1071 CAGCATATACCTGCGGATGTTCTCTCAGTTCAATTTTACCTGCTGCTTGAATCCG 1130  
Db 1182 CAGCATATACCTGCGGATGTTCTCTCAGTTCAATTTTACCTGCTGCTTGAATCCG 1241  
Qy 1131 AGCAATTCCTAATAAGGATTTTTCGAGCCCTTGTGACTATACAGTACAGTGTCTGA 1190  
Db 1242 AGCAATTCCTAATAAGGATTTTTCGAGCCCTTGTGACTATACAGTACAGTGTCTGA 1301  
Qy 1191 AAAGCGAGGCTACAGAGCTGTGCTAGTAAACAAGGGGAGCGGATTAATAGCCGCTTGT 1250  
|||||

Db 1302 AAAGCGAGGCTACAGAGACTGTGCTAGTAAACAAGGGGAGCGGATTAATAGCTTGT 1361  
Qy 1251 AGAGATGCTGTGCTGCTTGTGTGATCTCTCTCTGCTTCTCTTGTGCGACACC 1310  
Db 1362 AGAGATGCTGTGCTGCTTGTGTGATCTCTCTCTGCTTCTCTTGTGCGACACC 1421  
Qy 1311 ACTGATCTCTGCTGGAATCTTCTTAACTTAAACCAAGCAATTTCTGTGTCAAG 1370  
Db 1422 ACTGATCTCTGCTGGAATCTTCTTAACTTAAACCAAGCAATTTCTGTGTCAAG 1481  
Qy 1371 CTCAGTTTATTTCAACCCAGAAAGCTTCATTTTGTCTTCAATTTGGAATTTCTGT 1430  
Db 1482 CTCAGTTTATTTCAACCCAGAAAGCTTCATTTTGTCTTCAATTTGGAATTTCTGT 1541  
Qy 1431 TACTGCCAACAAGAGCTTCTGCGAAGGAGATATGACAGCTGCTGCTTGTGGT 1490  
Db 1542 TACTGCCAACAAGAGCTTCTGCGAAGGAGATATGACAGCTGCTGCTTGTGGT 1601  
Qy 1491 TGGTTCAGTTCTTCAACCCAAACATATGATGATCCATGAAAGACAGGGGGAAAGCCGTGGC 1550  
Db 1602 TGGTTCAGTTCTTCAACCCAAACATATGATGATCCATGAAAGACAGGGGGAAAGCCGTGGC 1661  
Qy 1551 TCCTAAGATATCCATCTCTCTGAAACAATAATCTTCTCACTTACAGATGACCCCTC 1610  
Db 1662 TCCTAAGATATCCATCTCTCTGAAACAATAATCTTCTCACTTACAGATGACCCCTC 1721  
Qy 1611 AATCCCATCATATATGCTGCTGCGAAGACCGGATAGCCCGCTTATTTGGTCTTACA 1670  
Db 1722 AATCCCATCATATATGCTGCTGCGAAGACCGGATAGCCCGCTTATTTGGTCTTACA 1781  
Qy 1671 AACTAGAGAAACCTCCAGAAACAACCCAGATGGAATTTTGGAGCAATGGTGGTT 1730  
Db 1782 AACTAGAGAAACCTCCAGAAACAACCCAGATGGAATTTTGGAGCAATGGTGGTT 1841  
Qy 1731 TTTTGGCTCAGGATATAGATAGGATTAATCTATTCAGAAAGAGCTGAGATTTCT 1790  
Db 1842 TTTTGGCTCAGGATATAGATAGGATTAATCTATTCAGAAAGAGCTGAGATTTCT 1901  
Qy 1791 TAAAGATGGATTTAATCTATCTAATAAGTTTCTCTCAAGATGCTCTGTGGGA 1850  
Db 1902 TAAAGATGGATTTAATCTATCTAATAAGTTTCTCTCAAGATGCTCTGTGGGA 1961  
Qy 1851 GGAGAAAGCCCGCAAGATATGATTAACAACAACATCCAGCTTCAAGGACAGAGTGGC 1910  
Db 1962 GGAGAAAGCCCGCAAGATATGATTAACAACAACATCCAGCTTCAAGGACAGAGTGGC 2021  
Qy 1911 GAGAATCTCTCTCAGAGAAACGGCATATTTATGTGTGTGAGATGCAAAAGATATGCG 1970  
Db 2022 RAGATCTCTCTCAGAGAAACGGCATATTTATGTGTGTGAGATGCAAAAGATATGCG 2081  
Qy 1971 CAAGATGTACATGATGCTCTTGTGCAATTAATTAAGCAAAAGCTTGAATAAACT 2030  
Db 2082 CAAGATGTACATGATGCTCTTGTGCAATTAATTAAGCAAAAGCTTGAATAAACT 2141  
Qy 2031 AGAAGCAATGAATAACCTGCGCATTTTAAAGAAAGAAACGCTACCTCAGGATATTTG 2090  
Db 2142 AGAAGCAATGAATAACCTGCGCATTTTAAAGAAAGAAACGCTACCTCAGGATATTTG 2201  
Qy 2091 GTCATAA 2097  
Db 2202 GTCATAA 2208  
|||||

RESULT 9  
US-10-029-386-6369  
; Sequence 6369, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO

FILE REFERENCE: A60MICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 34288  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 6369  
LENGTH: 591  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC008727.5  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.45  
OTHER INFORMATION: NT HIT: AF121205.1, EVALUE 0.00e+00  
OTHER INFORMATION: SWISSPROT HIT: Q9UPW0, EVALUE 3.30e+00  
OTHER INFORMATION: EST\_HUMAN HIT: A1132586.1, EVALUE 0.00e+00  
US-10-029-386-6369

Query Match 15.7%; Score 330; DB 16; Length 591;  
Best Local Similarity 99.7%; Pred. No. 1.3e-169;  
Matches 380; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 401 GTTTAGAACTTGTGTTGAGCCCTGATTTGCTGAGACTTGGCCAGCCCTCAGAAAGCATT 460  
DB 38 GTTTAGAACTTGTGTTGAGCCCTGATTTGCTGAGACTTGGCCAGCCCTCAGAAAGCATT 97  
QY 461 TTAGGTCAAGCAGAGCAAGAGAGATTAAGTGGGCACTCCCGGTGGATCAGCTGACAT 520  
DB 98 TTAGGTCAAGCAGAGCAAGAGAGATTAAGTGGGCACTCCCGGTGGATCAGCTGACAT 157  
QY 521 CCTTGAAGCAGAGCCTTGTGAAGTCAAGCTGCTACATTTGAATCTCAAGTGAAGCTTC 580  
DB 158 CCTTGAAGCAGAGCCTTGTGAAGTCAAGCTGCTACATTTGAATCTCAAGTGAAGCTTC 217  
QY 581 TGAATTCATGATTTCAAGAGAAAGATTTCTGAGGTTTGAAGCAAAATGCAAGTGAACA 640  
DB 218 TGAATTCATGATTTCAAGAGAAAGATTTCTGAGGTTTGAAGCAAAATGCAAGTGAACA 277  
QY 641 GGAACCAATCCAAATGTTGAATTAATGAAGACTTGAAGCTTCACTTACCCGTTGGTACCCC 700  
DB 278 GGAACCAATCCAAATGTTGAATTAATGAAGACTTGAAGCTTCACTTACCCGTTGGTACCCC 337  
QY 701 CACTCTCAAGCCTCTCTGAATATTCCTGTTTACCCCAAGATATTTTACAGGTACATC 760  
DB 338 CACTCTCAAGCCTCTCTGAATATTCCTGTTTACCCCAAGATATTTTACAGGTACATC 397  
QY 761 TGCAGAGTCTCTTGGCCAGG 781  
DB 398 TGCAGAGTCTCTTGGCCAGG 418

RESULT 10  
US-10-029-386-20100  
Sequence 20100, Application US/10029386  
Publication No. US20030194704A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
FILE REFERENCE: A60MICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 34288  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 20100  
LENGTH: 379  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC008727.5  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.45  
OTHER INFORMATION: NT HIT: g114729757, EVALUE 0.00e+00

OTHER INFORMATION: SWISSPROT HIT: Q9UPW0, EVALUE 1.80e+00  
OTHER INFORMATION: EST\_HUMAN HIT: A1132586.1, EVALUE 0.00e+00  
US-10-029-386-20100

Query Match 15.6%; Score 328; DB 16; Length 379;  
Best Local Similarity 99.7%; Pred. No. 1.6e-168;  
Matches 378; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 402 TTTAGAACTTGTGTTGAGCCCTGATTTGCTGAGACTTGGCCAGCCCTCAGAAAGCATT 461  
DB 1 TTTAGAACTTGTGTTGAGCCCTGATTTGCTGAGACTTGGCCAGCCCTCAGAAAGCATT 60  
QY 462 TAGGTCAAGCAGAGCAAGAGAGATTAAGTGGGCACTCCCGGTGGATCAGCTGACATC 521  
DB 61 TAGGTCAAGCAGAGCAAGAGAGATTAAGTGGGCACTCCCGGTGGATCAGCTGACATC 120  
QY 522 CTGAGAGCAGCCTTGTGAATTAATGAAGCTTGAAGCTTCACTTACCTCAAGTGAAGCTTC 581  
DB 121 CTGAGAGCAGCCTTGTGAATTAATGAAGCTTGAAGCTTCACTTACCTCAAGTGAAGCTTC 180  
QY 582 GAGATTCATGATTTCAAGAGAAAGATTTCTGAGGTTTGAAGCAAAATGCAAGTGAACAG 641  
DB 181 GAGATTCATGATTTCAAGAGAAAGATTTCTGAGGTTTGAAGCAAAATGCAAGTGAACAG 240  
QY 642 CAACCAATCCAAATGTTGAATTAATGAAGCTTGAAGCTTCACTTACCTCAAGTGAACAG 701  
DB 241 CAACCAATCCAAATGTTGAATTAATGAAGCTTGAAGCTTCACTTACCTCAAGTGAACAG 300  
QY 702 ACTCTCAAGCCTCTCTGAATATTCCTGTTTACCCCAAGATATTTTACAGGTACATC 761  
DB 301 ACTCTCAAGCCTCTCTGAATATTCCTGTTTACCCCAAGATATTTTACAGGTACATC 360  
QY 762 GCAGAGTCTCTTGGCCAG 780  
DB 361 GCAGAGTCTCTTGGCCAG 379

RESULT 11  
US-10-029-386-1735  
Sequence 1735, Application US/10029386  
Publication No. US20030194704A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
FILE REFERENCE: A60MICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 34288  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 1735  
LENGTH: 591  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC021609.3  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.4  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.8  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.2  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.2  
OTHER INFORMATION: NT HIT: AF121205.1, EVALUE 0.00e+00  
OTHER INFORMATION: EST\_HUMAN HIT: A1132586.1, EVALUE 0.00e+00  
OTHER INFORMATION: SWISSPROT HIT: Q9UPW0, EVALUE 3.30e+00  
US-10-029-386-1735

Query Match 13.3%; Score 279; DB 16; Length 591;  
Best Local Similarity 99.5%; Pred. No. 1.3e-141;  
Matches 379; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 401 GTTTAGAACTTGTGTGAGCCGTTGATTTGCTGAGCTTGGCCAGCCCTCAGAAAGCATT 460  
Db 38 GTTTAGAACTTGTGTGAGCCGTTGATTTGCTGAGCTTGGCCAGCCCTCAGAAAGCATT 97  
QY 461 TTAGGTCAAGCAGAGCAGAAAGAGATTAAGTGGCGCACTCCGGTGGCATCCTGCAT 520  
Db 98 TTAGGTCAAGCAGAGCAGAAAGAGATTAAGTGGCGCACTCCGGTGGCATCCTGCAT 157  
QY 521 CCTTGAGAGCAGACTTGTGAAGTCAAGCTGTACACATTTGAATCTCAAGTGCAGCTTC 580  
Db 158 CTTGAGAGCAGACTTGTGAAGTCAAGCTGTACACATTTGAATCTCAAGTGCAGCTTC 217  
QY 581 TGAGATTCAGATGATTCAGAAAGAAAGATTTCTGAGCTTTGAAAGCAAAATGCAGTGACA 640  
Db 218 TGAGATTCAGATGATTCAGAAAGAAAGATTTCTGAGCTTTGAAAGCAAAATGCAGTGACA 277  
QY 641 GCAACCAATCCAAATGTTGTAATTAAGATCTTGAGTCTGACATTAACCCGTGGTACCCC 700  
Db 278 GCAACCAATCCAAATGTTGTAATTAAGATCTTGAGTCTGACATTAACCCGTGGTACCCC 337  
QY 701 CACTCTCACAAGCCTCTCTGAATATTTCTGATTTACCCCAAGATATTTACAGGTACATC 760  
Db 338 CACTCTCACAAGCCTCTCTGAATATTTCTGATTTACCCCAAGATATTTACAGGTACATC 397  
QY 761 TGCAGAGTCTCTTGGCCAGG 781  
Db 398 TGCAGAGTCTCTTGGCCAGG 418

## RESULT 12

US-10-029-386-15435  
; Sequence 15435, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO  
; FILE REFERENCE: AEOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 15435  
; LENGTH: 379  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC021609.3  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.4  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.8  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.2  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.2  
; OTHER INFORMATION: SWISSPROT HIT: Q9UPW0, EVALUATE 1.80e+00  
; OTHER INFORMATION: EST HUMAN HIT: AU132586.1, EVALUATE 0.00e+00  
; OTHER INFORMATION: NT HIT: g14729757, EVALUATE 0.00e+00  
US-10-029-386-15435

Query Match 13.2%; Score 277; DB 16; Length 379;  
Best Local Similarity 99.5%; Pred. No. 1.5e-140;  
Matches 377; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 402 TTTAGAACTTGTGTGAGCCGTTGATTTGCTGAGCTTGGCCAGCCCTCAGAAAGCATT 461  
Db 1 TTTAGAACTTGTGTGAGCCGTTGATTTGCTGAGCTTGGCCAGCCCTCAGAAAGCATT 60  
QY 462 TAGTCAGAGCAGAGCAGAAAGAGATTAAGTGGCGCACTCCGGTGGCATCCTGCATC 521  
Db 61 TAGTCAGAGCAGAGCAGAAAGAGATTAAGTGGCGCACTCCGGTGGCATCCTGCATC 120

QY 522 CTTGAGAGCAGACTTGTGAAGTCAAGCTGTACACATTTGAATCTCAAGTGCAGCTTCT 581  
Db 121 CTTGAGAGCAGACTTGTGAAGTCAAGCTGTACACATTTGAATCTCAAGTGCAGCTTCT 180  
QY 582 GAAATTCAGATGATTCAGAAAGAAAGATTTCTGAGCTTTGAAAGCAAAATGCAGTGACA 641  
Db 181 GAAATTCAGATGATTCAGAAAGAAAGATTTCTGAGCTTTGAAAGCAAAATGCAGTGACA 240  
QY 642 CAACCAATCCAAATGTTGTAATTAAGATCTTGAGTCTGACATTAACCCGTGGTACCCC 701  
Db 241 CAACCAATCCAAATGTTGTAATTAAGATCTTGAGTCTGACATTAACCCGTGGTACCCC 300  
QY 702 ACTCTCACAAGCCTCTCTGAATATTTCTGATTTACCCCAAGATATTTACAGGTACATC 761  
Db 301 ACTCTCACAAGCCTCTCTGAATATTTCTGATTTACCCCAAGATATTTACAGGTACATC 360  
QY 762 GCAGAGTCTCTTGGCCAG 780  
Db 361 GCAGAGTCTCTTGGCCAG 379

## RESULT 13

US-10-741-600-17757  
; Sequence 17757, Application US/10741600  
; Publication No. US20050026169A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001499  
; CURRENT APPLICATION NUMBER: US/10/741,600  
; CURRENT FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 73997  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17757  
; LENGTH: 43985  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-741-600-17757

Query Match 12.7%; Score 266; DB 21; Length 43985;  
Best Local Similarity 99.5%; Pred. No. 2.1e-134;  
Matches 366; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 401 GTTTAGAACTTGTGTGAGCCGTTGATTTGCTGAGCTTGGCCAGCCCTCAGAAAGCATT 460  
Db 14836 GTTTAGAACTTGTGTGAGCCGTTGATTTGCTGAGCTTGGCCAGCCCTCAGAAAGCATT 14895  
QY 461 TTAGGTCAAGCAGAGCAGAAAGAGATTAAGTGGCGCACTCCGGTGGCATCCTGCAT 520  
Db 14896 TTAGGTCAAGCAGAGCAGAAAGAGATTAAGTGGCGCACTCCGGTGGCATCCTGCAT 14955  
QY 521 CTTTGAGAGCAGACTTGTGAAGTCAAGCTGTACACATTTGAATCTCAAGTGCAGCTTC 580  
Db 14956 CTTTGAGAGCAGACTTGTGAAGTCAAGCTGTACACATTTGAATCTCAAGTGCAGCTTC 15015  
QY 581 TGAGATTCAGATGATTCAGAAAGAAAGATTTCTGAGCTTTGAAAGCAAAATGCAGTGACA 640  
Db 15016 TGAGATTCAGATGATTCAGAAAGAAAGATTTCTGAGCTTTGAAAGCAAAATGCAGTGACA 15075  
QY 641 GCAACCAATCCAAATGTTGTAATTAAGATCTTGAGTCTGACATTAACCCGTGGTACCCC 700  
Db 15076 GCAACCAATCCAAATGTTGTAATTAAGATCTTGAGTCTGACATTAACCCGTGGTACCCC 15135  
QY 701 CACTCTCACAAGCCTCTCTGAATATTTCTGATTTACCCCAAGATATTTACAGGTACATC 760  
Db 15136 CACTCTCACAAGCCTCTCTGAATATTTCTGATTTACCCCAAGATATTTACAGGTACATC 15195  
QY 761 TGCAGAGG 768  
Db 15196 TGCAGAGG 15203

RESULT 14  
US-10-029-386-633/c  
Sequence 633, Application US/10029386  
Publication No. US20030194704A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: AEOMICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 34288  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 633  
LENGTH: 525  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC021609.3  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.48  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.58  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.52  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.57  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.79  
OTHER INFORMATION: SWISSPROT HIT: P37039, EVALU 1.00e-06  
OTHER INFORMATION: EST HUMAN HIT: BF346446.1, EVALU 1.00e-98  
OTHER INFORMATION: NT HIT: AF121212.1, EVALU 0.00e+00  
US-10-029-386-633

Query Match 9.0%; Score 188; DB 16; Length 525;  
Best Local Similarity 100.0%; Pred. No. 1.1e-91; Mismatches 0; Indels 0; Gaps 0;

Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1765 TTCAGAAAAGAGCTCAGACATTTCTTAAGCATGGATCTTAATCATCTAAAGTTTCC 1824  
DB 234 TTCAGAAAAGAGCTCAGACATTTCTTAAGCATGGATCTTAATCATCTAAAGTTTCC 175  
QY 1825 TTCTCAAGAGATGCTCTGTTGGGAGAGAAAGCCCGCAAGATGTACAGCAAC 1884  
DB 174 TTCTCAAGAGATGCTCTGTTGGGAGAGAAAGCCCGCAAGATGTACAGCAAC 115  
QY 1885 ATCCAGCTTCATGCGCAGACAGTGGGAGAAATCTCTCCAGAGAAAGCCCATATTAT 1944  
DB 114 ATCCAGCTTCATGCGCAGACAGTGGGAGAAATCTCTCCAGAGAAAGCCCATATTAT 55  
QY 1945 GTGTGTGG 1952  
DB 54 GTGTGTGG 47

RESULT 15  
US-10-029-386-14338/c  
Sequence 14338, Application US/10029386  
Publication No. US20030194704A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: AEOMICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 34288  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 14338  
LENGTH: 175  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC021609.3

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.48  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.58  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.52  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.57  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.79  
OTHER INFORMATION: SWISSPROT HIT: O61608, EVALU 4.00e-04  
OTHER INFORMATION: EST HUMAN HIT: AA085543.1, EVALU 7.00e-94  
OTHER INFORMATION: NT HIT: g113325067, EVALU 5.00e-94  
US-10-029-386-14338

Query Match 8.3%; Score 175; DB 16; Length 175;  
Best Local Similarity 100.0%; Pred. No. 1.5e-84; Mismatches 0; Indels 0; Gaps 0;

Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1770 AAAAGAGCTCAGACATTTCTTAAGCATGGATCTTAATCATCTAAAGTTTCCCTTC 1829  
DB 175 AAAAGAGCTCAGACATTTCTTAAGCATGGATCTTAATCATCTAAAGTTTCCCTTC 116  
QY 1830 AAGAGTGTCTCTGTTGGGAGAGAAAGCCCGCAAGATGTACAGCAACATCCA 1889  
DB 115 AAGAGTGTCTCTGTTGGGAGAGAAAGCCCGCAAGATGTACAGCAACATCCA 56  
QY 1890 GCTTCATGCGCAGACAGTGGGAGAAATCTCTCCAGAGAAAGCCCATATTAT 1944  
DB 55 GCTTCATGCGCAGACAGTGGGAGAAATCTCTCCAGAGAAAGCCCATATTAT 1

Search completed: August 27, 2005, 17:33:29  
Job time : 904.401 secs

**This Page Blank (uspio)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 23:18:31 ; Search time 4539.53 Seconds  
(without alignments)  
17558.328 Million cell updates/sec

Title: US-09-371-347A-45

Perfect score: 2094

Sequence: 1 atgaggaggttcctgtact.....ttcagatattgtgcataa 2094

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:\*  
1: gb\_esc1:\*  
2: gb\_esc2:\*  
3: gb\_hlc:\*  
4: gb\_esc3:\*  
5: gb\_esc4:\*  
6: gb\_esc5:\*  
7: gb\_esc6:\*  
8: gb\_esc81:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1349	64.4	3100	3	BC062577 Homo sapi
2	956	45.7	3143	3	BC035977 Homo sapi
3	719	34.3	908	5	BX348674 BX348674
4	689	32.9	874	4	BM801462 AGENCOURT
5	623	29.8	646	7	CN260357 170004241
6	565	27.0	565	1	AU279788 AU279788
7	543	25.9	877	1	AU124440 AU124440
8	531	25.4	1061	5	BQ218755 AGENCOURT
9	517	24.7	826	4	BI772430 603055786
10	512	24.5	776	6	CB997527 AGENCOURT
11	507	24.2	834	5	BU941078 AGENCOURT
12	485	21.7	822	1	AU132586 AU132586
13	448	21.4	591	2	AW965709 EST377782
14	446	21.3	818	6	CD559384 AGENCOURT
15	434	20.7	591	6	BI025283 RCS-MT025
16	431	20.6	974	5	BX375211 BX375211
17	406	19.4	710	5	BU570323 AGENCOURT
18	384	18.3	527	4	BI025277 RCS-MT025
19	374	17.9	579	7	CN260360 170006001
20	361	17.2	692	7	CN260359 170004706
21	360	17.2	685	4	BM049352 603626120
22	359	17.1	499	6	CD704108 EST20635
23	354	16.9	386	1	AA279726 EST2410.r
24	351	16.8	839	4	BG531787 602560355

25	343	16.4	852	5	BQ431497	BQ431497 AGENCOURT
26	341	16.3	395	4	BM838530	BM838530 K-EST0114
27	340	16.2	526	2	AW952883	AW952883 EST344953
28	337	16.1	818	7	CE995233	CE995233 AGENCOURT
29	335	16.0	413	2	BF810368	BF810368 RCS-C1041
30	335	16.0	413	2	BF810479	BF810479 RCS-C1014
31	332	15.9	366	1	AA085543	AA085543 ZN44h11.r
32	311	14.9	478	4	BM754488	BM754488 K-EST0031
33	302	14.4	440	4	CG877205	CG877205 OV3-HT046
34	297	14.2	416	6	CB996520	CB996520 AGENCOURT
35	292	13.9	528	2	BE301292	BE301292 ba89d07.k
36	291	13.9	664	7	CR768694	CR768694 DKF2P459K
37	291	13.9	667	7	CR770923	CR770923 DKF2P469N
38	291	13.9	767	7	CR557482	CR557482 DKF2P469K
39	282	13.5	521	6	CB164340	CB164340 K-EST0225
40	276	13.2	481	7	CR549172	CR549172 DKF2P459J
41	272	13.0	301	1	AL704780	AL704780 DKF2P686M
42	264	12.6	366	6	CB298361	CB298361 Z2019.te
43	257	12.3	366	2	BF808461	BF808461 QV1-C1017
44	257	12.3	368	1	AA355001	AA355001 EST63417
45	252	12.0	324	1	AA469901	AA469901 ZL94D04.r

#### ALIGNMENTS

RESULT 1	BC062577	3100 bp	mrna	linear	HTC 25-NOV-2003
LOCUS	Homo sapiens	cdna	clone IMAGE:5189058,	containing frame-shift errors.	
DEFINITION	BC062577.1	GI:36511756			
ACCESSION	BC062577				
VERSION	BC062577.1	GI:36511756			
KEYWORDS	HTC.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Strausberg,R.D., Fellingold,B.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Schmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Scheffer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,D., Hsieh,F., Diachenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Scapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Uedlin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hilyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,B.J., Lu,X., Gibbs,R.A., Fahy,J., Helton,E., Kettelman,M., Madan,A., Rodriques,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butlerfield,Y.S., Krzywicki,M.I., Skalska,U., Smallus,D.E., Butcher,A., Schein,J.E., Jones,S.J. and Marra,M.A.				
TITLE	human and mouse cDNA sequences				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	99 (26),	16899-16903	(2002)	
MEDLINE	2388257				
PUBMED	12477932				
REFERENCE	2 (bases 1 to 3100)				
AUTHORS	Strausberg,R.				
TITLE	Direct Submision				
JOURNAL	Submitted (24-NOV-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,				
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>				
COMMENT	Contact: MGC help desk Email: <a href="mailto:cgabs-r@mail.nih.gov">cgabs-r@mail.nih.gov</a> Tissue Procurement: Life Technologies, Inc.				

cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web Site: <http://www.nisc.nih.gov/>  
Contact: nisc\_mgc@hghri.nih.gov  
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,  
Diatch, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,  
Hansen, N., Ho, S.-L., Karlene, E., Kwong, P., Laric, P., Legaspi, R.,  
Maduro, Q.L., Maiello, C., Maskeri, B., Mastrian, S.D., McCluskey, J.C.,  
McDowell, J., Pearson, R., Stankitpop, S., Thomas, P.J., Touchman, J.W.,  
Tsurgren, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LNLN at: <http://image.lnl.gov>  
Series: IRAC Plate: 135 Row: e Column: 21  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 4505278  
This clone has the following problem: frame shifted.

## FEATURES

## source

1. 3100  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5189058"  
/issue\_type="Colon, Kidney, Stomach, adult, whole pooled"  
/clone\_lib="NIH\_MGC\_116"  
/lab\_host="DH10B"  
/note="Vector: PCMV-SPORT6"

## ORIGIN

Query Match 64.4%; Score 1349; DB 3; Length 3100;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1809; Conservative 0; Mismatches 2; Indels 4; Gaps 2;  
283 GGTCTGGTATTCAGAAATACACTACTTTGCAATGGGGGGAAGATATGATTAACGA 342  
172 GGTCTGGTATTCAGAAATACACTACTTTGCAATGGGGGGAAGATATGATTAACGA 231  
343 CTTCAAGAGCTTGGAGCCCGGCATTTCATGACATGAGACATGACATGACTGTAGT 402  
232 CTTCAAGAGCTTGGAGCCCGGCATTTCATGACATGAGACATGACATGACTGTAGT 291  
403 TTGGAATTGTTGGTGGCCGTGGATTGCTGAGCTCGGCGACGCCCAAGAAAGCATTT 462  
292 TTGGAATTGTTGGTGGCCGTGGATTGCTGAGCTCGGCGACGCCCAAGAAAGCATTT 351  
463 AGGTCAAGCAGAGCAGAGATTAAGTGGCGACTCCCGGTGGCATCACTGTATCC 522  
352 AGGTCAAGCAGAGCAGAGATTAAGTGGCGACTCCCGGTGGCATCACTGTATCC 411  
523 TTGAGACAGACCTTGTGAAGTCAAGAGCTGTACACTTGAATTCAGTTCAGCTTCTG 582  
412 TCGAGGACAGACCTTGTGAAGTCAAGAGCTGTACACTTGAATTCAGTTCAGCTTCTG 471  
583 AGATTGATGATTGAGGAAGAGATTCGAGGTTTGAAGCAAAATGAGTGAAGCAGC 642  
472 AGATTGATGATTGAGGAAGAGATTCGAGGTTTGAAGCAAAATGAGTGAAGCAGC 531  
643 AACCAATCAATGTTGTAATTTGAAGCTTGAAGTCTCACTTACCCGTTGATACCCCA 702  
532 AACCAATCAATGTTGTAATTTGAAGCTTGAAGTCTCACTTACCCGTTGATACCCCA 591  
703 CTCTCAAGCCTCTCTGAATATTCTGTGTTTACCCCAAGATATTTTACAGTATCATCTG 762  
592 CTCTCAAGCCTCTCTGAATATTCTGTGTTTACCCCAAGATATTTTACAGTATCATCTG 651  
763 CAGGAGCTCTTGGCCAGAGGAAGCAAGATATCTTGACTTCAGAGATTCAGTTT 822  
652 CAGGAGCTCTTGGCCAGAGGAAGCAAGATATCTTGACTTCAGAGATTCAGTTT 711

823 CAAGTCCAAATTCAAAGGACGTTCAACTTACTACGATGATGCCATTAAGCACTCTG 882  
712 CAAGTCCAAATTCAAAGGACGTTCAACTTACTACGATGATGCCATTAAGCACTCTG 771  
883 CTGTGTAATTTGACATTTTCAAAATACAGCTTTCTTATCAGCTTGGAGTGGCTTACG 942  
772 CTGTGTAATTTGACATTTTCAAAATACAGCTTTCTTATCAGCTTGGAGTGGCTTACG 831  
943 GTGATGCTCCCTTACAGTATTTGAGTACAAAGGCTTCTCCAAAGCTGACGTTGAA 1002  
832 GTGATGCTCCCTTACAGTATTTGAGTACAAAGGCTTCTCCAAAGCTGACGTTGAA 891  
1003 GATTAAGAGAGCAGCTGCGCTTTTGAATTAAGGACAGACAAAGAAAGAGAGCT 1062  
892 GATTAAGAGAGCAGCTGCGCTTTTGAATTAAGGACAGACAAAGAAAGAGAGCT 950  
1063 ACCTTACCCAGACATATCTGCGGAGATTTCTCTCCAGATTCATTTTACCTGCTTCT 1122  
951 ACCTTACCCAGACATATCTGCGGAGATTTCTCTCCAGATTCATTTTACCTGCTTCT 1010  
1123 GAAATCCGAGCAATTCCTTAAGAAAGCAATTTTGCAGAGCCCTTGTGACTATACAGTGAC 1182  
1011 GAAATCCGAGCAATTCCTTAAGAAAGCAATTTTGCAGAGCCCTTGTGACTATACAGTGAC 1070  
1183 AGTCTGAAAGGCGAGGCTTACAGAGCTGTGAGTAAACAAGGGCAGCCGATTATAGC 1242  
1071 AGTCTGAAAGGCGAGGCTTACAGAGCTGTGAGTAAACAAGGGCAGCCGATTATAGC 1130  
1243 CGCTTTGTACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1302  
1131 CGCTTTGTACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1190  
1303 CAGCCACCACTAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1362  
1191 CAGCCACCACTAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1250  
1363 TGTGAAAGCTCAAGTTTATTTTACCCAGAAAGCTCCATTTGTCTTCAACATTTGTGAA 1422  
1251 TGTGAAAGCTCAAGTTTATTTTACCCAGAAAGCTCCATTTGTCTTCAACATTTGTGAA 1310  
1423 TTTCTGTCTTACGCTTCAACAAGAGGTTTCCGGAAGAGGATATGACAGCTGCTGCGC 1482  
1311 TTTCTGTCTTACGCTTCAACAAGAGGTTTCCGGAAGAGGATATGACAGCTGCTGCGC 1370  
1483 TTGTTGTTGCTTCAAGTTCTTCAAGCAAAATATCATGATCCCATGAAAGCAGGGAAA 1542  
1371 TTGTTGTTGCTTCAAGTTCTTCAAGCAAAATATCATGATCCCATGAAAGCAGGGAAA 1430  
1543 GCCCTGCTCTTAAGATATCATCTCTCTGGAACAACAAATCTTTTCACTTACAGAT 1602  
1431 GCCCTGCTCTTAAGATATCATCTCTCTGGAACAACAAATCTTTTCACTTACAGAT 1490  
1603 GACCCCTCAATCCCAATCAATTAAGTGGGTCCAGGAACCGGCAATAGCCCGTTATTTGG 1662  
1491 GACCCCTCAATCCCAATCAATTAAGTGGGTCCAGGAACCGGCAATAGCCCGTTATTTGG 1550  
1663 TTCTTCAACATATGAGAGAACTTCAAGAACCAACCCAGATGAAATTTTGAAGCAATG 1722  
1551 TTCTTCAACATATGAGAGAACTTCAAGAACCAACCCAGATGAAATTTTGAAGCAATG 1610  
1723 TG---GTTTTTGGCTGAGGCAATTAAGATAGGATTTATCTATTCAGAAAAGGCTCAGA 1779  
1611 TGTTGTTTGGCTGAGGCAATTAAGATAGGATTTATCTATTCAGAAAAGGCTCAGA 1670  
1780 CATTTCTTAAGATAGGATTTTAACTCATCTTAAAGTTTCTTCTCAAGATAGCTCCT 1839  
1671 CATTTCTTAAGATAGGATTTTAACTCATCTTAAAGTTTCTTCTCAAGATAGCTCCT 1730  
1840 GTTGGGAGAGAGAACCCAGCAAGATATGTAAGAACCAACATCCAGTTTCAAGGCGAG 1899  
1731 GTTGGGAGAGAGAACCCAGCAAGATATGTAAGAACCAACATCCAGTTTCAAGGCGAG 1790

QY 1900 CAGGTGGCGAGATCTCTCCAGAGAAACGGCATATTATGTGTGAGATGCAAG 1959  
DB 1791 CAGGTGGCGAGATCTCTCCAGAGAAACGGCATATTATGTGTGAGATGCAAG 1850  
QY 1960 AATATGGCCAGAGATGATATGATGCTTGTGCAATATATAGCAAGAGTTGAGTT 2019  
DB 1851 AATATGGCCAGAGATGATATGATGCTTGTGCAATATATAGCAAGAGTTGAGTT 1910  
QY 2020 GAAAACTAGAGCAATGAAAAACCTGGCACCCTTTAAAAAGAGAAAAACGTTACCTTCAG 2079  
DB 1911 GAAAACTAGAGCAATGAAAAACCTGGCACCCTTTAAAAAGAGAAAAACGTTACCTTCAG 1970  
QY 2080 GATATTTGGTCATAA 2094  
DB 1971 GATATTTGGTCATAA 1985

RESULT 2  
LOCUS BC035977 3143 bp mRNA linear HTC 20-SEP-2002  
DEFINITION Homo sapiens, clone IMAGE:4611253, mRNA.  
ACCESSION BC035977  
VERSION BC035977.1 GI:23243305  
KEYWORDS HTC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 3143)  
Strausberg, R.  
Direct Submission  
Submitted (31-Jul-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
Contact: MGC help desk  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: CLONTECH  
CDNA Library Preparation: CLONTECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www-shgc.stanford.edu>  
Contract: (Dickson, Mark) [mcdepxil@stanford.edu](mailto:mcdepxil@stanford.edu)  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
R. M.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
Series: IRML Plate: 41 Row: g Column: 2  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 4505278  
This clone has the following problem: frame shifted.  
FEATURES  
source  
1..3143  
location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4611253"  
/tissue\_type="Kidney"  
/clone\_id="NIH\_MGC\_75"  
/lab\_host="DH10B"  
/note="Vector: pDNR-L1B"

ORIGIN  
Query Match 45.7%; Score 956; DB 3; Length 3143;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1056; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGAGAGTTCTCTTACTATATGTACACAGCAGGCAAGGCAATCGCAGAA 60  
DB 52 ATGAGAGAGTTCTCTTACTATATGTACACAGCAGGCAAGGCAATCGCAGAA 111

QY 61 GAAATGTGTAGCAGCTGTGTGATCATGATATTTCTGCAGATCTTCACTGATATGATGA 120  
DB 112 GAAATGTGTAGCAGCTGTGTGATCATGATATTTCTGCAGATCTTCACTGATATGATGA 171  
QY 121 TCCGATAGATATGACCTTAAAAACGAAACAGCTCTCTTGTGTGTGTGTTCTACACG 180  
DB 172 TCCGATAGATATGACCTTAAAAACGAAACAGCTCTCTTGTGTGTGTGTTCTACACG 231  
QY 181 GGCACCGAGACCCACCCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACA 240  
DB 232 GGCACCGAGACCCACCCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACA 291  
QY 241 CTGCCGTTGATTTCTTGTCTCACTGGGATATGGGTTACTGGGTTCTGGTATTCAGAA 300  
DB 292 CTGCCGTTGATTTCTTGTCTCACTGGGATATGGGTTACTGGGTTCTGGTATTCAGAA 351  
QY 301 TACACCTTACTTTGCAATGAGGAGGAAATATTTGATTAACGACTTCAAGAGCTTGAGCC 360  
DB 352 TACACCTTACTTTGCAATGAGGAGGAAATATTTGATTAACGACTTCAAGAGCTTGAGCC 411  
QY 361 CGGCATTTCTATGACATCTGACATGACAGATGACCTGTGATGTTTGAAGCTTGTGAG 420  
DB 412 CGGCATTTCTATGACATCTGACATGACAGATGACCTGTGATGTTTGAAGCTTGTGAG 471  
QY 421 CCGTGATTTGCTGAACTCTGGCAGACCCCTCAGAAAGCAATTTTAAAGTCAAGAGACAA 480  
DB 472 CCGTGATTTGCTGAACTCTGGCAGACCCCTCAGAAAGCAATTTTAAAGTCAAGAGACAA 531  
QY 481 GAGAGATATAGTGGGCACTCCCGGTGACATCCCTGATCCTTGAAGACAGACCTTGTG 540  
DB 532 GAGAGATATAGTGGGCACTCCCGGTGACATCCCTGATCCTTGAAGACAGACCTTGTG 591  
QY 541 AAGTCAGAGCTGCTACATGATTAATCTCAAGTCAAGCTTCTGAAGTTGATGATTCAGGA 600  
DB 592 AAGTCAGAGCTGCTACATGATTAATCTCAAGTCAAGCTTCTGAAGTTGATGATTCAGGA 651  
QY 601 AGAAAGATTCGAGTTTGAAGCAAAATGACATGAGCAAGCAACCAATCCATGTTGTA 660  
DB 652 AGAAAGATTCGAGTTTGAAGCAAAATGACATGAGCAAGCAACCAATCCATGTTGTA 711  
QY 661 ATTGAAGACTTGTAGTCTCACTTACCCGTTGATACCCCACTCTCAAGACCTTCTGTG 720  
DB 712 ATTGAAGACTTGTAGTCTCACTTACCCGTTGATACCCCACTCTCAAGACCTTCTGTG 771  
QY 721 AATATTCCTGTTTACCCCGAGATATTTTACAGTATCTGACAGAGTCTTGGCCAG 780  
DB 772 AATATTCCTGTTTACCCCGAGATATTTTACAGTATCTGACAGAGTCTTGGCCAG 831  
QY 781 GAGGAAGCGAGATCTGTGACTTCAGAGATTCAGTTTTCAGAGTGCATTTCAAG 840  
DB 832 GAGGAAGCGAGATCTGTGACTTCAGAGATTCAGTTTTCAGAGTGCATTTCAAG 891  
QY 841 GAGTTCAACTTACTACGATATGATGCAATTAACCACTCTGCTGTGATGATTCAGAT 900  
DB 892 GAGTTCAACTTACTACGATATGATGCAATTAACCACTCTGCTGTGATGATTCAGAT 951  
QY 901 TCAATATACAGACTTTTCTATCAGCTGTGAGATGCTTCAAGCTTATGCTTAAACAT 960  
DB 952 TCAATATACAGACTTTTCTATCAGCTGTGAGATGCTTCAAGCTTATGCTTAAACAT 1011  
QY 961 GATTTCTGAGTACCAAGCTTATCCAAAGACCTGAGCTTGAAGATTAAGAGACCTGC 1020  
DB 1012 GATTTCTGAGTACCAAGCTTATCCAAAGACCTGAGCTTGAAGATTAAGAGACCTGC 1071  
QY 1021 GTCTTTTGAATTAAGGACAGACAAAGAAAGG 1058  
DB 1072 GTCTTTTGAATTAAGGACAGACAAAGAAAGG 1109

RESULT 3  
LOCUS BX348674 908 bp mRNA linear EST 08-APR-2004  
BX348674

DEFINITION BX348674 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
CDNA clone CS00C010Y11 5-PRIME, mRNA sequence.  
ACCESSION BX348674  
VERSION BX348674.1 GI:30375301  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE Li W.B., Gruber C., Jesse J. and Polyes D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 Evry cedex - FRANCE  
Email: seqref@genoscope.cns.fr Web: www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
3392.f  
For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?c=CS0BAG0062B02\_CS00490\_1&c=3392.f

FEATURES  
source  
Location/Qualifiers  
1..908  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DC010Y11"  
/issue\_type="NEUROBLASTOMA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."  
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN  
Query Match 34.3%; Score 719; DB 5; Length 908;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 719; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 674 AGTCCTCACTTACCCGTTCCGTAACCCCACTCTCACAAAGCCTCTGATATTCCTGTT 723  
DB 28 AGTCCTCACTTACCCGTTCCGTAACCCCACTCTCACAAAGCCTCTGATATTCCTGTT 87  
QY 734 TACCCCCGAATTTTACAGTACATCTGCAGAGAGTCTCTTGCCAGAGAGAAAGCCAG 793  
DB 88 TACCCCCGAATTTTACAGTACATCTGCAGAGAGTCTCTTGCCAGAGAGAAAGCCAG 147  
QY 794 TATCTGTGACTTCAGCAGATCCAGTTTTCAGAGCCAAATTTCAAGAGGAGTTCAACTTA 853  
DB 148 TATCTGTGACTTCAGCAGATCCAGTTTTCAGAGCCAAATTTCAAGAGGAGTTCAACTTA 207  
QY 854 CTACGATGATGCGATTAACCACTCTGCTGTGATGATGACATTTCAATACAGACT 913  
DB 208 CTACGATGATGCGATTAACCACTCTGCTGTGATGATGACATTTCAATACAGACT 267  
QY 914 TTTCTATCAGCTGAGATGCTTCAAGCGTATCTGCCCTTAACAGTATTCGAGGTAC 973  
DB 268 TTTCTATCAGCTGAGATGCTTCAAGCGTATCTGCCCTTAACAGTATTCGAGGTAC 327  
QY 974 AAGGCCATCTCCAAAGCTCAGCTTGAAGTAAAGAGAGCTGCTCTTTGAAA 1033  
DB 328 AAGGCCATCTCCAAAGCTCAGCTTGAAGTAAAGAGAGCTGCTCTTTGAAA 387  
QY 1034 TAAAGCAGACACAAAGAGAGAGCTTACCCAGCATATACCTGCGGAGTGT 1093  
DB 388 TAAAGCAGACACAAAGAGAGAGCTTACCCAGCATATACCTGCGGAGTGT 447  
QY 1094 CTCTCAGTTCAATTTTACCTGCTGTGAATTCGAGCAATTCCTAAAGGCAATTTT 1153

DB 448 CTCTCAGTTCAATTTTACCTGCTGTGAATTCGAGCAATTCCTAAAGGCAATTTT 507  
QY 1154 TCGAGACCCCTTGTGACATATACAGTACAGTGTGAAAAGCGAGGCTACAGAGCTGT 1213  
DB 508 TCGAGACCCCTTGTGACATATACAGTACAGTGTGAAAAGCGAGGCTACAGAGCTGT 567  
QY 1214 GCAGTAAACAGAGGCGAGCCGATTAATGCGCTTTGTACAGATGCTGTGCTGCTGT 1273  
DB 568 GCAGTAAACAGAGGCGAGCCGATTAATGCGCTTTGTACAGATGCTGTGCTGCTGT 627  
QY 1274 TGGATCTCTCTCTGCTTCCCTTCTTGTGACGACCACTCACTGCTCCGCTGGAATC 1333  
DB 628 TGGATCTCTCTCTGCTTCCCTTCTTGTGACGACCACTCACTGCTCCGCTGGAATC 687  
QY 1334 TTCTTAACCTTCAACCCAGACCATATTCGCTGTGCAAGCTCAAGTTATTTTCAACCCAGA 1392  
DB 688 TTCTTAACCTTCAACCCAGACCATATTCGCTGTGCAAGCTCAAGTTATTTTCAACCCAGA 746

RESULT 4  
BM801462 874 bp mRNA linear EST 05-MAR-2002  
LOCUS BM801462  
DEFINITION AGENCOURT\_6459212 NIH\_MGC\_88 Homo sapiens cDNA clone IMAGE:5560477  
5', mRNA sequence.  
ACCESSION BM801462  
VERSION BM801462.1 GI:19118285  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.  
1 (bases 1 to 874)  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM12286 row: 1 column: 14  
High quality sequence stop: 710.

FEATURES  
source  
Location/Qualifiers  
1..874  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5560477"  
/issue\_type="duodenal adenocarcinoma, cell line"  
/lab\_note="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC 88"  
/note="Organ: small intestine; Vector: pCMV-SPORT6;  
Site 1: NotI; Site 2: SalI; Cloned unidirectionally;  
oligo-dT primed. Average insert size 1.767 kb. Library  
enriched for full-length clones and constructed by Life  
Technologies. Note: this is a NIH\_MGC Library."

ORIGIN  
Query Match 32.9%; Score 689; DB 4; Length 874;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 739; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ATGAGAGGTTCTGTACTATATGCTACACAGAGGAGAGGCAAGCCATGCGAGAA 60  
DB 50 ATGAGAGGTTCTGTACTATATGCTACACAGAGGAGAGGCAAGCCATGCGAGAA 109  
QY 61 GAAATGTGAGGAGAGCTGTGTATGATGATTTTCTGCAAGTCTTCACTGATATTAATGAA 120

```

Db      110 GAAATATGTGACCAAGCTGTGTGATCATGATTTCTGCAGATCTTCACTGTATTAGTGA 169
Qy      121 TCGATTAAGTATGACCTTAAACCCGAAAGAGCTCCCTGTGTGTGTGTCTTCAACAG 180
Db      170 TCCGAAAGTATGACCTTAAACCCGAAAGAGCTCCCTGTGTGTGTGTCTTCAACAG 229
Qy      181 GGCACCGGAGACCCACCCGACACAGCCGCAAGTTGTGTAGAAATACAGAACCAACA 240
Db      230 GGCACCGGAGACCCACCCGACACAGCCGCAAGTTGTGTAGAAATACAGAACCAACA 289
Qy      241 CTGCGGTGATTTCTTCTCTCACTGCGGTATGAGTTACTGAGTCTCGGTATTCAGA 300
Db      290 CTGCGGTGATTTCTTCTCTCACTGCGGTATGAGTTACTGAGTCTCGGTATTCAGA 349
Qy      301 TACACTTACTTTTGCAATGGGGGAAATATGATTAACGATTCAGAGCTTGAGCC 360
Db      350 TACACTTACTTTTGCAATGGGGGAAATATGATTAACGATTCAGAGCTTGAGCC 409
Qy      361 CGGCAATTTCTATGACACTGACATGACATGACTGTGTAGATTGAACTTGTTGAG 420
Db      410 CGGCAATTTCTATGACACTGACATGACATGACTGTGTAGATTGAACTTGTTGAG 469
Qy      421 CGGTGATTTCTGACACTGACCTGACCTGACAGCAATTTAGTCAAGCAGAGACA 480
Db      470 CGGTGATTTCTGACACTGACCTGACCTGACAGCAATTTAGTCAAGCAGAGACA 529
Qy      481 GAGGATTAAGTGTGGCCTATCCCGTGTGATCATCTGATCTTGAAGACACCTTGTG 540
Db      530 GAGGATTAAGTGTGGCCTATCCCGTGTGATCATCTGATCTTGAAGACACCTTGTG 589
Qy      541 AAGTCAGAGCTGTACACATTTGATCTCAAGTCAAGCTTCTGAGATTTCAGATTCAGA 600
Db      590 AAGTCAGAGCTGTACACATTTGATCTCAAGTCAAGCTTCTGAGATTTCAGATTCAGA 649
Qy      601 AGAAAGATTTCTGAGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATCCAATGTTGA 660
Db      650 AGAAAGATTTCTGAGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATCCAATGTTGA 709
Qy      661 ATTGAAGACTTTGAGTCTTCACTTACCCGTTGGTACCCCACTCTCAAGCTTCTG 720
Db      710 ATTGAAGACTTTGAGTCTTCACTTACCCGTTGGTACCCCACTCTCAAGCTTCTG 769
Qy      721 AATATTCCTGTTTACCCCC 740
Db      770 AATATTCCTGTTTACCCCC 789

```

```

RESULT 5
CN260357      646 bp      mRNA      linear      EST 16-MAY-2004
DEFINITION    17000424179730 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
ACCESSION     CN260357
VERSION       CN260357.1 GI:4726771
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
               1 (bases 1 to 646)
               Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fiek, G.J.,
               Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
               Lebowitz, J. and Stanton, L.W.
               Transcriptome characterization elucidates signaling networks that
               control human ES cell growth and differentiation
               Nat. Biotechnol. 22 (6), 707-716 (2004)
               Contact: Brandenberger R
               Regenerative Medicine
               Geron Corporation
               230 Constitution Drive, Menlo Park, CA 94025, USA
               Tel: 650 473 8658
               Fax: 650 473 7760
               Email: rbrandenberger@geron.com
               Insert Length: 646      Std Error: 0.00.

```

## FEATURES

source

Location/Qualifiers

1..646

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/tissue\_type="embryonic stem cells, cell lines H1, H7, and

H9"

/clone\_id="GRN ES"

/note="oligo dt primed, full-length enriched cDNA library

from undifferentiated hbs cell lines H1 (p32), H7 (p29),

and H9 (p26) maintained in feeder-free conditions"

## ORIGIN

Query Match 29.8%; Score 623; DB 7; Length 646;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 623; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      987 MAGCTGCAGCTTGAAGATTAAGAGAGACATGCGTCTTTGAAATTAAGGACAGAC 1046
Db      24 MAGCTGCAGCTTGAAGATTAAGAGAGACATGCGTCTTTGAAATTAAGGACAGAC 83
Qy      1047 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1106
Db      84 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 143
Qy      1107 TTTTACCTGCTGTCTTGAAGATTCAGAGCAATTCCTAATAAGGAGATTTTTCAGAGCCCTGT 1166
Db      144 TTTTACCTGCTGTCTTGAAGATTCAGAGCAATTCCTAATAAGGAGATTTTTCAGAGCCCTGT 203
Qy      1167 GGAATTAACAGTGAAGTGTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1226
Db      204 GGAATTAACAGTGAAGTGTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 263
Qy      1227 GGCAGCCGATTAAGCCGCTTTGTACAGATGCTGTGCTCTGTGTGATCTCTCTCT 1286
Db      264 GGCAGCCGATTAAGCCGCTTTGTACAGATGCTGTGCTCTGTGTGATCTCTCTCTCT 323
Qy      1287 GCGTTTCCCTCTTTCGACGACACACAGTCACTGCTCGCTGAGAAATCTTCTTAACTTCA 1346
Db      324 GCGTTTCCCTCTTTCGACGACACACAGTCACTGCTCGCTGAGAAATCTTCTTAACTTCA 383
Qy      1347 ACCAGAGACCAATTTGTGTGACAGTCAAGTTATTTTCAACAGAGAAAGCTCAATTTGT 1406
Db      384 ACCAGAGACCAATTTGTGTGACAGTCAAGTTATTTTCAACAGAGAAAGCTCAATTTGT 443
Qy      1407 CTTCAACATTTGTGAATTTCTGTCTACTGCAACACAGAGTTTTCGCGAAGGAGATAG 1466
Db      444 CTTCAACATTTGTGAATTTCTGTCTACTGCAACACAGAGTTTTCGCGAAGGAGATAG 503
Qy      1467 TACAGGCTGGCTGGCTGTGTGTGCTTCACTTCTTCAAGCAACATACATGATCCCA 1526
Db      504 TACAGGCTGGCTGGCTGTGTGTGCTTCACTTCTTCAAGCAACATACATGATCCCA 563
Qy      1527 TGAAGACAGCGGAGAAAGCCCTGCTCTAAGATTCATCTCTCTGAGAACCAATTC 1586
Db      564 TGAAGACAGCGGAGAAAGCCCTGCTCTAAGATTCATCTCTCTGAGAACCAATTC 623
Qy      1587 TTTCACTTACCAATGACCCCT 1609
Db      624 TTTCACTTACCAATGACCCCT 646

```

## RESULT 6

A0279788

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

```

A0279788      565 bp      mRNA      linear      EST 31-JUL-2003
LOCUS         A0279788 CHONS2 Homo sapiens cDNA clone CHONS2001448 5', mRNA
DEFINITION    sequence.
ACCESSION     A0279788
VERSION       A0279788.1 GI:28299015
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

REFERENCE 1 (bases 1 to 565)  
AUTHORS Imabayashi, H., Mori, T., Gojo, S., Kiyono, T., Sugiyama, T., Irie, R.,  
Isogai, T., Hata, J., Tomoya, Y., and Umezawa, A.  
TITLE Redifferentiation of dedifferentiated chondrocytes and  
chondrogenesis of human bone marrow stromal cells via chondrosphere  
formation with expression profiling by large-scale cDNA analysis  
JOURNAL Exp. Cell Res. 288 (1), 35-50 (2003)  
MEDLINE 22760698  
PUBMED 12878157  
COMMENT Contact: Takao Isogai  
Genomics Laboratory  
Helix Research Institute  
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: genomics@hri.co.jp  
HRI human cDNA project, Sugiyama, T.; Wakamatsu, A.; Irie, R.;  
Umezawa, A.; Fukuma, M.; Kusakari, S.; Hata, J.; Ishii, S.; Yamamoto, J.;  
Iseno, Y.; Saito, K.; Nakamura, Y.; Masuno, Y.; Nagai, K.; Isogai, T.  
HRI human cDNA project; cDNA library construction & 5'-end one  
pass sequencing: Helix Research Institute.  
location/Qualifiers  
1..565  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CHONS2001448"  
/cell\_type="chondrocytes"  
/clone\_1ib="CHONS2"  
/note="Vector: PME18SFL3"

ORIGIN  
Query Match 27.0%; Score 565; DB 1; Length 565;  
Best Local Similarity 100.0%; Pred. No. 2.1e-298;  
Matches 565; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

819 TTTTCAAGTCCCAATTTCAAAAGGCACTTCACTTCAAGATGATGCCATAAAACCAC 878  
1 TTTTCAAGTCCCAATTTCAAAAGGCACTTCACTTCAAGATGATGCCATAAAACCAC 60

879 TCTGCTGGTGAATTTGAGCACTTCAAAATACAGACTTTTCTATCAGCTGAGATGCTT 938  
61 TCTGCTGGTGAATTTGAGCACTTCAAAATACAGACTTTTCTATCAGCTGAGATGCTT 120

939 CAGCTGATCTGCTTCAAGATGATCTGAGTACAAAGCTTCAAGAGCTGAGCT 998  
121 CAGCTGATCTGCTTCAAGATGATCTGAGTACAAAGCTTCAAGAGCTGAGCT 180

999 TGAAGATTAAG 1058  
181 TGAAGATTAAG 240

1059 AGCTACCTTACCCAGAGATTAATCTGCGGAGATGCTCTCAGTTCACTTTTCTCTG 1118  
241 AGCTACCTTACCCAGAGATTAATCTGCGGAGATGCTCTCAGTTCACTTTTCTCTG 300

1119 TCTTGAATCCAGCAATTTCTTAAAGGCAATTTTTCGAGCCCTTGTGACTATACAG 1178  
301 TCTTGAATCCAGCAATTTCTTAAAGGCAATTTTTCGAGCCCTTGTGACTATACAG 360

1179 TGAAGTGTGAAAAGCGAGGCTACAGAGCTGTGAGTAAACAGGGGAGAGCCGATTA 1238  
361 TGAAGTGTGAAAAGCGAGGCTACAGAGCTGTGAGTAAACAGGGGAGAGCCGATTA 420

1239 TAGCGCGTTTGTGAGATGCTGTGCTGCTTGTGAGATCTCTCTGCTTCTCCCTTC 1298  
421 TAGCGCGTTTGTGAGATGCTGTGCTGCTTGTGAGATCTCTCTGCTTCTCCCTTC 480

1299 TTGCGAGCAGCACTCAGTCTCCTGCTCGAATCTTCTTAACTTCAACCGAGACATA 1358  
481 TTGCGAGCAGCACTCAGTCTCCTGCTCGAATCTTCTTAACTTCAACCGAGACATA 540

ORIGIN  
Query Match 25.9%; Score 543; DB 1; Length 877;  
Best Local Similarity 100.0%; Pred. No. 2.8e-286;  
Matches 543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1143 AAAGGATTTTTCGAGCCCTTGTGACTATACAGTGAACAGTGTGAAAAGGCGAGGCT 1202  
192 AAAGGATTTTTCGAGCCCTTGTGACTATACAGTGAACAGTGTGAAAAGGCGAGGCT 251

1203 ACAAGAGCTGTGAGTAAACAAGGGGAGCCGATTAATAGCCGCTTGTGAAGATGCTG 1262  
252 ACAAGAGCTGTGAGTAAACAAGGGGAGCCGATTAATAGCCGCTTGTGAAGATGCTG 311

1263 TGGCTGCTTGTGAGATCTCTCTGCTTCCCTTCTTGGCCAGCCAGCTCAGTCTCCT 1322  
312 TGGCTGCTTGTGAGATCTCTCTGCTTCCCTTCTTGGCCAGCCAGCTCAGTCTCCT 371

1323 GCTCGAAGATCTTCCCTTCAACCTTCAACCGAGACATATGTGTGGAAGCTCAAGTTATT 1382  
372 GCTCGAAGATCTTCCCTTCAACCTTCAACCGAGACATATGTGTGGAAGCTCAAGTTATT 431

1383 TCACCGAGGAAGAGCTCATTTTGTCTTCAACATTTGTGAATTTCTGTACTGCGCAAC 1442  
432 TCACCGAGGAAGAGCTCATTTTGTCTTCAACATTTGTGAATTTCTGTACTGCGCAAC 491



Db	Accession	Version	Source	Organism	Reference Authors Title Journal Comment
Db	61				
QY	1115				
Db	121				
QY	1175				
Db	181				
QY	1235				
Db	241				
QY	1295				
Db	301				
QY	1355				
Db	361				
QY	1415				
Db	421				
QY	1475				
Db	481				
RESULT 9					
LOCUS					
DEFINITION					
ACCESSION					
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
COMMENT					
FEATURES					
Source					



## ORIGIN

range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH\_MGC Library."

Query Match 24.7%; Score 517; DB 4; Length 826;  
Best Local Similarity 99.7%; Pred. No. 5.7e-272;  
Matches 617; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATGAGAGGTTTCTGTACTATATGCTACACAGAGGAGACAGCAAGCCATCGCAGAA 60
DB 53 ATGAGAGGTTTCTGTACTATATGCTACACAGAGGAGACAGCAAGCCATCGCAGAA 112
QY 61 GAATGCTGTAGCAAGCTGTGTATCATGATTTTCTCAGATCTTCACTGTATTAGTAA 120
DB 113 GAATATATGTAGCAAGCTGTGTATCATGATTTTCTCAGATCTTCACTGTATTAGTAA 172
QY 121 TCCGATTAAGTATGACCTTAATAACCGAAAGCTCTCTGTGTGTGTGTCTTACACAG 180
DB 173 TCCGATTAAGTATGACCTTAATAACCGAAAGCTCTCTGTGTGTGTGTCTTACACAG 232
QY 181 GGACCGGAGACCCACCCGACACAGCCCGAAGTTTGTAAAGAAATACAGAAACAAACA 240
DB 233 GGACCGGAGACCCACCCGACACAGCCCGAAGTTTGTAAAGAAATACAGAAACAAACA 292
QY 241 CTGCGCGTTGATTTCTTGTCTCAGCTCGGTATGAGTTTCTGCGTCTCGGTATTCAGAA 300
DB 293 CTGCGCGTTGATTTCTTGTCTCAGCTCGGTATGAGTTTCTGCGTCTCGGTATTCAGAA 352
QY 301 TACACCTACTTTTGCATATGGGGGGAAGATTAATTGATTAAGACTTCAAGAGCTTGGAGCC 360
DB 353 TACACCTACTTTTGCATATGGGGGGAAGATTAATTGATTAAGACTTCAAGAGCTTGGAGCC 412
QY 361 CGGCAATTTCTATGACACTGTGACATGACATGCTGTGTAGATTTAGAACTTGTGTGAG 420
DB 413 CGGCAATTTCTATGACACTGTGACATGACATGCTGTGTAGATTTAGAACTTGTGTGAG 472
QY 421 CCGTGATTTCTGTGACCTGTGCGCACGCTTCAAGAAAGATTTTATGCTACAGACAGACAA 480
DB 473 CCGTGATTTCTGTGACCTGTGCGCACGCTTCAAGAAAGATTTTATGCTACAGACAGACAA 532
QY 481 GAGAGATTAAGTGGCGGACCTCCGGTGGACATCCCTGACCTTGAAGAGACAGACTTGTG 540
DB 533 GAGAGATTAAGTGGCGGACCTCCGGTGGACATCCCTGACCTTGAAGAGACAGACTTGTG 592
QY 541 AAGTCAGAGCTGTACACATTTGAATCTCAAGTCGAGCTTCTGAGATTGATTCAGAGA 600
DB 593 AAGTCAGAGCTGTACACATTTGAATCTCAAGTCGAGCTTCTGAGATTGATTCAGAGA 652
QY 601 AGAAGGATTTCTGAGTTT 619
DB 653 AGAAGGATTTCTGAGTTT 671
```

RESULT 10  
CB97527 776 bp mRNA linear EST 01-MAY-2003  
LOCUS  
DEFINITION  
AGENCOURT 13620640 NIH MGC 148 Homo sapiens cDNA clone  
CB97527  
ACCESSION  
CB97527 GI:30292047  
VERSION  
KEYWORDS  
EST.  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
1 (bases 1 to 776)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)

Tissue Procurement: Dr. Stefan Hanson  
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help  
and advice from Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.lnl.gov>  
Plate: NDAM365 row: 1 column: 21  
High quality sequence stop: 564.  
Location/Qualifiers  
1. 776  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3038684"  
/cissue\_type="pre-eclmptic placenta"  
/lab\_host="DH10B Tona"  
/clone\_idb="NIH\_MGC\_148"  
/note="Organ: Placenta; Vector: pBluescript; Site\_1:  
all-XhoI; Site\_2: BamHI; Library is oligo-dT primed and  
directionally cloned using primer  
5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert  
size 2.3 kb and normalized to ROT 5. This is a primary  
library enriched for full-length clones and constructed  
using the Cap-trapper method (Carninci, in preparation).  
Library constructed by M. Brownstein (NIH/NHGRI,  
National Institutes of Health). Note: this is a NIH\_MGC  
Library."

FEATURES  
source

## ORIGIN

Query Match 24.5%; Score 512; DB 6; Length 776;  
Best Local Similarity 99.7%; Pred. No. 3.2e-269;  
Matches 612; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATGAGAGGTTTCTGTACTATATGCTACACAGAGGAGACAGCAAGCCATCGCAGAA 60
DB 88 ATGAGAGGTTTCTGTACTATATGCTACACAGAGGAGACAGCAAGCCATCGCAGAA 147
QY 61 GAATGCTGTAGCAAGCTGTGTATCATGATTTTCTCAGATCTTCACTGTATTAGTAA 120
DB 148 GAATATATGTAGCAAGCTGTGTATCATGATTTTCTCAGATCTTCACTGTATTAGTAA 207
QY 121 TCCGATTAAGTATGACCTTAATAACCGAAAGCTCTCTGTGTGTGTGTCTTACACAG 180
DB 208 TCCGATTAAGTATGACCTTAATAACCGAAAGCTCTCTGTGTGTGTGTCTTACACAG 267
QY 181 GGACCGGAGACCCACCCGACACAGCCCGAAGTTTGTAAAGAAATACAGAAACAAACA 240
DB 268 GGACCGGAGACCCACCCGACACAGCCCGAAGTTTGTAAAGAAATACAGAAACAAACA 327
QY 241 CTGCGCGTTGATTTCTTGTCTCAGCTCGGTATGAGTTTCTGCGTCTCGGTATTCAGAA 300
DB 328 CTGCGCGTTGATTTCTTGTCTCAGCTCGGTATGAGTTTCTGCGTCTCGGTATTCAGAA 387
QY 301 TACACCTACTTTTGCATATGGGGGGAAGATTAATTGATTAAGACTTCAAGAGCTTGGAGCC 360
DB 388 TACACCTACTTTTGCATATGGGGGGAAGATTAATTGATTAAGACTTCAAGAGCTTGGAGCC 447
QY 361 CGGCAATTTCTATGACACTGTGACATGACATGCTGTGTAGATTTAGAACTTGTGTGAG 420
DB 448 CGGCAATTTCTATGACACTGTGACATGACATGCTGTGTAGATTTAGAACTTGTGTGAG 507
QY 421 CCGTGATTTCTGTGACCTGTGCGCACGCTTCAAGAAAGATTTTATGCTCAAGAGACAA 480
DB 508 CCGTGATTTCTGTGACCTGTGCGCACGCTTCAAGAAAGATTTTATGCTCAAGAGACAA 567
QY 481 GAGAGATTAAGTGGCGGACCTCCGGTGGACATCCCTGACCTTCTGAGACAGACTTGTG 540
DB 568 GAGAGATTAAGTGGCGGACCTCCGGTGGACATCCCTGACCTTCTGAGACAGACTTGTG 627
QY 541 AAGTCAGAGCTGTACACATTTGAATCTCAAGTCGAGCTTCTGAGATTGATTCAGAGA 600
```

```

Db      628  AAGTCAGAGCTGTACACATTGAATCTCAATCGAGCTTTGAGATTGATGATTCAGGA 687
Oy      601  AGAAGGATTTCTGA 614
        |||||
Db      688  AGAAGGATTTCTGA 701

RESULT 11
LOCUS   BU941078
DEFINITION
AGENCOURT_10540067 NIH MGC 128 Homo sapiens cDNA clone
IMAGE:6712893 5', mRNA sequence.
BU941078
BU941078.1  GI:24129897
EST.

Homo sapiens (human)
Homo sapiens
Mammarytota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 834)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: NCI
CDNA Library Preparation: Michael Brownstein Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.lnl.gov
Plate: LNCM302 row: e column: 21
High quality sequence stop: 586.

```

FEATURES	Location/Qualifiers
source	1. .834

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6712893"
/tissue_type="mixed (pool of 40 RNAs)"
/lab_host="DH10B (T1-phage-resistant)"
/clone_id="NIH MGC 127"
/notes="Vector: pDNR-LIB; Site_1: SfiI (ggccattagcgc);
Site_2: SfiI (ggccgcctcgcc); Double-stranded cDNA was
prepared from a pool of 40 cell line polyA+ RNAs (bladder
- 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -
4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,
kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell -
5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%,
salivary gland - 1.3%, and skin - 2.3%). 5' and 3'
adaptors were used in cloning as follows:
5'-AAGCAGTGTGATCAACAGAGATGGCCATTACGGCCGGG-3' and
5'-ATTTCAGAGCGCGAGCGCGCGGACATC-dT(30)NN-3'. Full-length
enriched library was constructed using the Clontech
Creator SMART kit and size-selected to contain the >2 kb
size fraction (other fractions present in NIH MGC 126 and
NIH MGC 127). Library created in the laboratory of T.
Usdin, M.D., Ph.D. (NIH, NIH). Note: this is a NIH_MGC
library."

```

	Query Match	Similarity	Score	DB	Length	Best Local	Pred.	No.	Matches	Conservative	Mismatches	Indels	Gaps
QY	379	GGAGATGAGATGACCTGCTGCTTAAACTGTGTGTGAGCCCTGATTCGTGACATC	24.2%	507	834	99.7%	1.8e-266	607	0	2	0	0	
DB	3	GGAGATGAGATGACTGTGTGTGATTAACCTGTGTGTGACCCCTGATTCGTGACATC											
QY	439	TGGCCAGCCTTCAGGAACATTTTATGATCAAGCAGAGCAGAGAGATTAATGTGCCCA			498								
DB	63	TGGCCAGCCTTCAGGAACATTTTATGATCAAGCAGAGCAGAGAGATTAATGTGCCCA			122								

Oy	499	CTCCGGGTGGATCACTGATATCCCTTGAAGACAGACTTGGAACTGACAGCTGAC	558
Db	123	CTCCGGGTGGATCACTGATATCCCTTGAAGACAGACTTGGAACTGACAGCTGAC	182
Oy	559	ATTGAATCTCAAGTCAGCTTCTGAGATTGATGATTCAGGAAGAAAGATTCGAGTT	618
Db	183	ATTGAATCTCAAGTCAGCTTCTGAGATTGATGATTCAGGAAGAAAGATTCGAGTT	242
Oy	619	TTGAAGCAAAATGCAATGAAACAGAACCAATCCATGTGTAATYGAAGACTTGAATCC	678
Db	243	TTGAAGCAAAATGCAATGAAACAGAACCAATCCATGTGTAATYGAAGACTTGAATCC	302
Oy	679	TCACCTTACCGGTGGTACCCCACTTCAACAAGCTCTCTGAATATTTCTGGTTTACC	738
Db	303	TCACCTTACCGGTGGTACCCCACTTCAACAAGCTCTCTGAATATTTCTGGTTTACC	362
Oy	739	CCGAATATTTACAGTACATCTGACAGAGTCTTGGCCAGAGGAAAGCCAAATGATCT	798
Db	363	CCGAATATTTACAGTACATCTGACAGAGTCTTGGCCAGAGGAAAGCCAAATGATCT	422
Oy	799	GTGACTTCAGCAGATCCAGTTTTTCAAGTGCCAATTTCAAGGCAAGTTCAACTACTACG	858
Db	423	GTGACTTCAGCAGATCCAGTTTTTCAAGTGCCAATTTCAAGGCAAGTTCAACTACTACG	482
Oy	859	AATGATGCCAATAAAACCACTCTGCTGGTGAATTTGACATTTCAATAACAGACTTTTCC	918
Db	483	AATGATGCCAATAAAACCACTCTGCTGGTGAATTTGACATTTCAATAACAGACTTTTCC	542
Oy	919	TATCAGCTTGAAGATGCTTCAAGGTGATCTGCCCTTAACAGTCAATTTGAGGTCAAAAGC	978
Db	543	TATCAGCTTGAAGATGCTTCAAGGTGATCTGCCCTTAACAGTCAATTTGAGGTCAAAAGC	602
Oy	979	CTACTCCAA 987	
Db	603	CTACTCCAA 611	

RESULT	12
AU132586	
LOCUS	
DEFINITION	AU132586 822 bp mRNA linear EST 01-AUG-2002
ACCESSION	AU132586 NT2RP4 Homo sapiens cDNA clone NT2RP4000141 5', mRNA sequence.
VERSION	AU132586
KEYWORDS	AU132586.1 GI:10992940
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 822) Oca,T., Sugiyama,T., Ishii,S., Suzuki,Y., Saito,K., Yamamoto,U., Nishikawa,T., Nakamura,Y., Nagai,T., Sungano,S., Masuno.Y. and Isogai,T. HRI human cDNA project (Oca,T., Sugiyama,T., Ishii,S., Suzuki,Y., Saito,K., Yamamoto,J., Nishikawa,T., Nakamura,Y., Nagai.T., Sugano.S., Masuno.Y., Isogai,T.) Unpublished (2000) Contact: Takao Isogai
JOURNAL	
COMMENT	Contact: Takao Isogai

```

FEATURES
source
    Helix Research Institute
    1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
    Tel.: 81-438-52-3975
    Fax: 81-438-52-3986
    Email: genomcshri.co.jp
    HRI human cDNA project: 5'- & 3'-end one pass sequencing: Helix
    Research Institute; cDNA library construction: Department of
    Virology, Institute of Medical Science, University of Tokyo, and
    Helix Research Institute.
    Location/Qualifiers
        1..822
        /organism="Homo sapiens"
        /mol_type="mRNA"

```

/db xref="taxon:9606"  
/clone="NT2RP4000141"  
/cell\_type="teratocarcinoma"  
/cell\_line="NT2"  
/clone\_1ib="NT2RP4"  
/note="Vector: PME18SFL3; mRNA from NT2 neuronal precursor  
cells after 2-weeks retinoic acid (RA) induction"

## ORIGIN

Query Match 21.7%; Score 455; DB 1; Length 822;  
Best Local Similarity 99.6%; Pred. No. 7.3e-238;  
Matches 555; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 283 GGTCTGGTATTCAGAAATACCTTACTTTGCAATGGGGGAAAGATTAATGTAACGA 342  
DB 182 GGTCTGGTATTCAGAAATACCTTACTTTGCAATGGGGGAAAGATTAATGTAACGA 241  
QY 343 CTTCAGAGCTTGAAGCCGGCAATTTCTATGACATGCAATGACATGCTGTAGGT 402  
DB 242 CTTCAGAGCTTGAAGCCGGCAATTTCTATGACATGCAATGACATGCTGTAGGT 301  
QY 403 TTGAACTTGTGTTGAGCCGTGATTCGACCTTGCCAGCCCTCAGAAAGCATTTT 462  
DB 302 TTGAACTTGTGTTGAGCCGTGATTCGACCTTGCCAGCCCTCAGAAAGCATTTT 361  
QY 463 AGCTCAGAGAGAGAGAGAGATTAAGTGGGCACTCCCGGTGGCATCTGATCC 522  
DB 362 AGGTCAAGCAGAGAGAGAGATTAAGTGGGCACTCCCGGTGGCATCTGATCC 421  
QY 523 TTGAGAGCAGAGCTTGAAGTCAAGCTCTACACATTAATCTCAAGTCAAGCTTCTG 582  
DB 422 TCGAGAGCAGAGCTTGAAGTCAAGCTCTACACATTAATCTCAAGTCAAGCTTCTG 481  
QY 583 AGATTGAGATTCAGAGAAAGATTCGAGGTTTGAAGCAAAATGCAATGAAACAGC 642  
DB 482 AGATTGAGATTCAGAGAAAGATTCGAGGTTTGAAGCAAAATGCAATGAAACAGC 541  
QY 643 AACCAATCCAAATTTGAAGATTTGAAGATTTGAAGATTTGAAGATTTGAAGATTTGA 702  
DB 542 AACCAATCCAAATTTGAAGATTTGAAGATTTGAAGATTTGAAGATTTGAAGATTTGA 601  
QY 703 CTCTCAAGAGCTCTGATATTTCTGTTTACCCCGAGATTAATTTTACAGTATCATCTG 762  
DB 602 CTCTCAAGAGCTCTGATATTTCTGTTTACCCCGAGATTAATTTTACAGTATCATCTG 661  
QY 763 CAGAGAGCTCTGAGCAGAGAGAGAGAGATTAATTTTACAGATTCAGATTTT 822  
DB 662 CAGAGAGCTCTGAGCAGAGAGAGAGAGATTAATTTTACAGATTCAGATTTT 721  
QY 823 CAGAGAGCTCTGAGCAGAGAGAGAGAGATTAATTTTACAGATTCAGATTTT 839  
DB 722 CAGAGAGCTCTGAGCAGAGAGAGAGAGATTAATTTTACAGATTCAGATTTT 738

RESULT 13  
AM965709 591 bp mRNA linear EST 01-JUN-2000  
LOCUS EST377782 MAGE resequences, MAGI Homo sapiens cDNA, mRNA sequence.  
DEFINITION AM965709  
ACCESSION AM965709.1 GI:8155545  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C.,  
Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and  
Quackenbush, J.  
Assessment of gene expression patterns in a model of colon tumor  
metastasis using a 19,200 element cDNA microarray  
Unpublished (2000)  
Contact: John Quackenbush

The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 3528  
Fax: 301 838 0208  
Email: johnd@tigr.org  
Plate: 218  
Seq primer: Reverse.  
Location/Qualifiers

## FEATURES

source

1..591  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_1ib="MAGE resequences, MAGI"  
/note="Vector: pBluescriptSkm"

## ORIGIN

Query Match 21.4%; Score 448; DB 2; Length 591;  
Best Local Similarity 100.0%; Pred. No. 4.9e-234;  
Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1192 AAGCGAGGCTACAGAGCTGTGACATAAACAAGGGCAGCCGATTAAGCCGTTTGA 1251  
DB 1 AAGCGAGGCTACAGAGCTGTGACATAAACAAGGGCAGCCGATTAAGCCGTTTGA 60  
QY 1252 CGAGATGCTGTGCTGCTGTTGATCTCTCCCTGCTTCCCTTCCAGCCAGCA 1311  
DB 61 CGAGATGCTGTGCTGCTGTTGATCTCTCCCTGCTTCCCTTCCAGCCAGCA 120  
QY 1312 CTCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1371  
DB 121 CTCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
QY 1372 TCGAGTTATTTCAACCGAGAAAGCTCAATTTGCTTCAATGTGAATTTCTGCT 1431  
DB 181 TCGAGTTATTTCAACCGAGAAAGCTCAATTTGCTTCAATGTGAATTTCTGCT 240  
QY 1432 ACTGCCAACAAGAGTTCTGGGAGAGAGATTAAGTGAAGGCTGCTGCTGCTGCTGCT 1491  
DB 241 ACTGCCAACAAGAGTTCTGGGAGAGAGATTAAGTGAAGGCTGCTGCTGCTGCTGCT 300  
QY 1492 GCTTCAGTTCTTCAGCAACAATACATGATCCATGAGAGACAGCGGAAAGCCCTGCT 1551  
DB 301 GCTTCAGTTCTTCAGCAACAATACATGATCCATGAGAGACAGCGGAAAGCCCTGCT 360  
QY 1552 CCTAAGATTCATCTCTCTGAGACACAAATTTCTTCACTTACAGATGACCCCTCA 1611  
DB 361 CCTAAGATTCATCTCTCTGAGACACAAATTTCTTCACTTACAGATGACCCCTCA 420  
QY 1612 ATCCCATCATTAATGAGTGGTCCAGGAA 1639  
DB 421 ATCCCATCATTAATGAGTGGTCCAGGAA 448

RESULT 14  
CD559384 818 bp mRNA linear EST 11-JUN-2003  
LOCUS CD559384  
DEFINITION AGENCOURT\_14401607 NIH\_MGC\_191 Homo sapiens cDNA clone  
ACCESSION IMAGC:30409775 5', mRNA sequence.  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
NIH-MGC http://mgc.ncl.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: c9apbs-r@mail.nih.gov  
 Tissue Procurement: Narayan Bhat  
 CDNA Library Preparation: Clontech Laboratories, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: NDCM198 row: n column: 24  
 High quality sequence stop: 484.

## FEATURES

source

1. .818  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30409775"  
 /tissue\_type="Pooled"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_1ib="NIH MGC 191"  
 /note="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggc);  
 Site 2: SfiI (ggcgccctggcc); Library is oligo-dT primed  
 and directionally cloned. PBMC - Peripheral Blood  
 Mononuclear Cells. RNA was pooled from 3/6hour stimulation  
 with PMA and ionomycin. 5' and 3' adaptors were used in  
 cloning as follows: 5' adaptor sequence:  
 5'-CACGCCATTATGACC-3' and 3' adaptor sequence:  
 5'-ATTCTAGAGCCGAGCGCCGACATG-RT(30)BR-3' (where B = A,  
 C, or G and N = A, C, G, or T). Average insert size 1.69  
 kb (range 0.70-5.0 kb). 15/15 colonies contained inserts  
 by PCR. This library was enriched for full-length clones  
 and was constructed by Clontech Laboratories (Palo Alto,  
 CA). Note: this is a NIH MGC Library."

## ORIGIN

Query Match 21.3%; Score 446; DB 6; Length 818;  
 Best Local Similarity 99.6%; Pred. No. 6.5e-233;  
 Matches 546; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

425 GGATTGCTGAGCTCTGGCCAGCCCTTGAAGAAAGATTGTTAGTCAAGACAGACAGAG 484  
 3 GGATTGCTGAGCTCTGGCCAGCCCTTGAAGAAAGATTGTTAGTCAAGACAGAGAG 62

485 AGATTAGTGGCGACTCCCGGTGGATGACCTGATCCTTGAAGACAGACAGCTTGAAGT 544  
 63 AGATTAGTGGCGACTCCCGGTGGATGACCTGATCCTTGAAGACAGACAGCTTGAAGT 122

545 CAGAGCTGTACACATGATGATCTCAAGTCAAGCTTCTGATTCATGATTCAGAGAA 604  
 123 CAGAGCTGTACACATGATGATCTCAAGTCAAGCTTCTGATTCATGATTCAGAGAA 182

605 AGGATTCTGAGGTTTGAAGCAAAATGATGAGACAGACCAATCAATGTTGTAATTG 664  
 183 AGGATTCTGAGGTTTGAAGCAAAATGATGAGACAGACCAATCAATGTTGTAATTG 242

665 AAGACTTGAAGTCCATCTTACCCGTTGGTACCCCACTCCAGCAAGAGCTCTGAAGA 724  
 243 AAGACTTGAAGTCCATCTTACCCGTTGGTACCCCACTCCAGCAAGAGCTCTGAAGA 302

725 TTCCTGTTTACCCCGAGATATTTACAGTACATCTGAGAGAGCTCTTGGCAGAGAG 784  
 303 TTCCTGTTTACCCCGAGATATTTACAGTACATCTGAGAGAGCTCTTGGCAGAGAG 362

785 AAAGCCAGATCTGTGACTTCAAGATCCAGTTTTCAGAGTCCCAATTTCAAGGCGAG 844  
 363 AAAGCCAGATCTGTGACTTCAAGATCCAGTTTTCAGAGTCCCAATTTCAAGGCGAG 422

845 TTCAACTTCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 904  
 423 TTCAACTTCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 482

905 ATACAGACTTTTCTATCAGCTGAGATGCTTCAAGCTGATCTGCTTAAAGATGAT 964  
 483 ATACAGACTTTTCTATCAGCTGAGATGCTTCAAGCTGATCTGCTTAAAGATGAT 542

QY 965 CTGAGCTA 972  
 DB 543 CTGAGCTA 550

## RESULT 15

BI025283 591 bp mRNA linear EST 14-JUN-2001  
 LOCUS BI025283  
 DEFINITION RCS-MT0259-020201-021-G04 MT0259 Homo sapiens CDNA, mRNA sequence.  
 ACCESSION BI025283  
 VERSION BI025283.1 GI:14431913  
 KEYWORDS EST.  
 SOURCE  
 ORGANISM Homo sapiens (human)

REFERENCE  
 AUTHORS Buxaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 591)  
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
 Nagai, M.A., da Silva, M. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bata, G.S., Simpson, D.H.,  
 Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V.,  
 O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
 Simpson, A.J.

## TITLE

Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 MEDLINE 20202663  
 PUBMED 10737800

## COMMENT

Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the PABSP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (<http://www.ludwig.org.br/scripte/gethtml2.pl?cl=RCS&c2=RCS-MT0259-020201-021-G04&t3=2001-02-02&t4=1>)  
 Seq primer: puc 18 forward  
 High quality sequence start: 78  
 High quality sequence stop: 590.  
 Location/Qualifiers

## FEATURES

source

1. .591  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_1ib="MT0259"  
 /note="Organ: marrow; Vector: puc18; Site 1: SmaI; Site 2:  
 SmaI; A mini-library was made by cloning products derived  
 from ORESTES PCR (U.S. Letters Patent application No.  
 196,716 - Ludwig Institute for Cancer Research) profiles  
 into the puc 18 vector. Reverse transcription of tissue  
 mRNA and cDNA amplification were performed under low  
 stringency conditions."

## ORIGIN

Query Match 20.7%; Score 434; DB 4; Length 591;  
 Best Local Similarity 99.8%; Pred. No. 2.5e-226;  
 Matches 484; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

525 GAGGACAGACTTTGAAATGACAGCTGTACACATTTGATCTCAAGTCAAGCTTCTGAG 584  
 102 GAGGACAGACTTTGAAATGACAGCTGTACACATTTGATCTCAAGTCAAGCTTCTGAG 161

585 ATTGCATGATTGAGAGAGAGAGATTTGAGAGTCTTTGAAGCAAAATGATGATGATGAT 644  
 162 ATTGCATGATTGAGAGAGAGATTTGAGAGTCTTTGAAGCAAAATGATGATGATGAT 221

645 CCAATCAATGTTGTAATGAAAGACTTGAAGTCTTACCTTACCGTGGTACCCCACT 704

```
Db      222 CCAATCCAAATGTTGTAATTAAGACTTGAAGTCTCACTTACCCGTTGGTACCCCACT 281
Qy      705 CTCACAAGCCTCTCTGAATATTCCTGTTTACCCCGAATATTTACAGGTACATCTGCA 764
Db      282 CTCACAAGCCTCTCTGAATATTCCTGTTTACCCCGAATATTTACAGGTACATCTGCA 341
Qy      765 GGAGTCTCTTGGCCAGAGGAAAGCCAGATCTGTGACTTCAGCAATCAATTCTTTCA 824
Db      342 GGAGACTCTTGGCCAGAGGAAAGCCAGATCTGTGACTTCAGCAATCAATTCTTTCA 401
Qy      825 AGTGCCAATTTCAAAGCAGTTCACTTACTAGCAATGATGCCATAAAACCACTGCGT 884
Db      402 AGTGCCAATTTCAAAGCAGTTCACTTACTAGCAATGATGCCATAAAACCACTGCGT 461
Qy      885 GGTAGAATTGACATTTCAATACAGACTTTTCTATCAGCCTGAGATGCCCTTCAGCGT 944
Db      462 GGTAGAATTGACATTTCAATACAGACTTTTCTATCAGCCTGAGATGCCCTTCAGCGT 521
Qy      945 GATCTGCCCTTAACAGTGATTTCTGAGTACAAAGCCTACTCCAAAGAAGCTGAGCTTGAAGA 1004
Db      522 GATCTGCCCTTAACAGTGATTTCTGAGTACAAAGCCTACTCCAAAGAAGCTGAGCTTGAAGA 581
Qy      1005 TAAAA 1009
Db      582 TAAAA 586
```

Search completed: August 27, 2005, 15:58:37  
Job time : 4541.53 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Comugen Ltd.

OW nucleic - nucleic search, using sw model

Run on: August 26, 2005, 19:17:21 / Search time 5998.62 Seconds  
(without alignments)  
16914.771 Million cell updates/sec

Title: US-09-371-347A-45  
Perfect score: 2094  
Sequence: 1 atgagagagttctgtctact.....ttcagatattgtgcataa 2094

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenBml:\*

1: gb\_ba:\*\n2: gb\_hic:\*\n3: gb\_in:\*\n4: gb\_om:\*\n5: gb\_ov:\*\n6: gb\_pac:\*\n7: gb\_ph:\*\n8: gb\_pl:\*\n9: gb\_pr:\*\n10: gb\_ro:\*\n11: gb\_scs:\*\n12: gb\_sy:\*\n13: gb\_un:\*\n14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1854	88.5	3259	6	AR144976 Sequence
2	1854	88.5	3259	6	AX050463 Sequence
3	1854	88.5	3259	9	AF025794 Homo sapi
4	1854	88.5	3291	9	AF121214 Homo sapi
5	1752	83.7	3241	6	CQ726091 Sequence
6	1701	81.2	3310	9	BC054816 Homo sapi
7	1220	58.3	2933	11	BV177620 Bv177620 sqm95800
8	1220	58.3	2933	11	BV178010 Bv178010 sqm97986
9	386	18.4	330	6	BD077780 5' EST of
10	381	18.2	1353	9	AF121205 Homo sapi
11	330	15.8	109626	9	AC010346 Homo sapi
12	330	15.8	110756	9	AC025174 Homo sapi
13	279	13.3	158199	2	AC022921 Homo sapi
14	279	13.3	167237	2	AC021609 Homo sapi
15	279	13.3	177596	2	AC091945 Homo sapi
16	189	9.0	1156	9	AF121210 Homo sapi
17	188	9.0	1034	9	AF121212 Homo sapi
18	188	9.0	167237	2	AC021609 Homo sapi
19	188	9.0	177596	2	AC091945 Homo sapi

20	183	8.7	1432	9	FI21202807	AF121208 Homo sapi
21	161	7.7	158199	2	AC022921	AC022921 Homo sapi
22	158	7.5	1256	9	FI21202802	AF121203 Homo sapi
23	158	7.5	2475	6	AR454615	AR454615 Sequence
24	158	7.5	2475	6	AX375651	AX375651 Sequence
25	155	7.4	2011	9	FI21202806	AF121207 Homo sapi
26	146	7.0	2214	9	FI21202812	AF121213 Homo sapi
27	129	6.2	4506	9	FI21202801	AF121206 Homo sapi
28	125	6.0	969	9	FI21202805	AF121206 Homo sapi
29	121	5.8	1119	9	FI21202810	AF121211 Homo sapi
30	119	5.7	1200	9	FI21202803	AF121204 Homo sapi
31	78	3.7	244	6	BD204284	BD204284 5' EST and
32	63	3.0	63	6	AX611833	AX611833 Sequence
33	60	2.9	60	6	CO539377	CO539377 Sequence
34	54	2.6	54	6	AX611839	AX611839 Sequence
35	54	2.6	54	6	AX611843	AX611843 Sequence
36	51	2.4	51	6	AX162161	AX162161 Sequence
37	48	2.3	48	6	AX611835	AX611835 Sequence
38	48	2.3	48	6	AX611841	AX611841 Sequence
39	47	2.2	183	6	CO670532	CO670532 Sequence
40	44	2.1	650	9	FI21202808	AF121209 Homo sapi
41	41	2.0	41	6	AX611845	AX611845 Sequence
42	38	1.8	38	6	AX611837	AX611837 Sequence
43	38	1.8	63	6	AX611834	AX611834 Sequence
44	32	1.5	238720	2	AC095949	AC095949 Rattus no
45	32	1.5	271339	2	AC131637	AC131637 Rattus no

## ALIGNMENTS

RESULT 1	AR144976	3259 bp	DNA	linear	PAT 08-AUG-2001
LOCUS	AR144976	Sequence 23 from patent US 6210950.			
DEFINITION	AR144976				
ACCESSION	AR144976.1	GI:15106843			
VERSION					
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 3259)				
AUTHORS	Johnson, W.G. and Stemroos, R.Scott.				
TITLE	Methods for diagnosing, preventing, and treating developmental disorders due to a combination of genetic and environmental factors				
JOURNAL	Patent: US 6210950-A 23 03-APR-2001;				
FEATURES	Location/Qualifiers				
source	1..3259				
ORIGIN	/organism="unknown"				
	/mol_type="unassigned DNA"				
Query Match	88.5%;	Score 1854;	DB 6;	Length 3259;	
Best Local Similarity	99.9%;	Pred. No. 0;			
Matches 2094;	Conservative	0;	Mismatches	0;	Indels 3; Gaps 1;
QY	1	ATGAGGAGGTTTCTGTACTATATGCTACACAGCAGGACGCAAGCCATCGAGAA	60		
DB	80	ATGAGGAGGTTTCTGTACTATATGCTACACAGCAGGACGCAAGCCATCGAGAA	139		
QY	61	GAATGTGACAGCTGTGATCATGATTTTCTGAGATCTTCACTGATTAATGA	120		
DB	140	GAATGTGACAGCTGTGATCATGATTTTCTGAGATCTTCACTGATTAATGA	199		
QY	121	TCGATTAAGTATGACCTTAACAAACGAAACAGCTCTTGTGTGTGTTCTACACG	180		
DB	200	TCGATTAAGTATGACCTTAACAAACGAAACAGCTCTTGTGTGTGTTCTACACG	259		
QY	181	GGACCGGAGACCCACCCGACAGCCGCGCAAGTTTGAAGAAATACGAACCAACA	240		
DB	260	GGACCGGAGACCCACCCGACAGCCGCGCAAGTTTGAAGAAATACGAACCAACA	319		
QY	241	CTGCCGTTGATTTCTTCTCACTGCGGTATGGGTTACTGGGTCCTCGATTCAGAA	300		

|||||  
Db 320 CTGCGGTTGATTTCTTGTCTCACTCGGATGAGTTTACGCGTCTCGGATTCAGAA 379  
QY 301 TACACCTACTTTTSCAATGGGGGAGATTAATTGATAAGCACTTCAAGAGCTTGAAGC 360  
Db 380 TACACCTACTTTTSCAATGGGGGAGATTAATTGATAAGCACTTCAAGAGCTTGAAGC 439  
QY 361 CGGCAATTTCTATGACACTGACATGACATGACATGACATGACATGACATGACATGAC 420  
Db 440 CGGCAATTTCTATGACACTGACATGACATGACATGACATGACATGACATGACATGAC 499  
QY 421 CCGGATTTGCTGATCTCTGGCCGAGCCCTCAGAAAGCACTTTAGGTCAAGAGAGCA 480  
Db 500 CCGGATTTGCTGATCTCTGGCCGAGCCCTCAGAAAGCACTTTAGGTCAAGAGAGCA 559  
QY 481 GAGGAGATTAAGTGGCGCACTCCCGGTGACATCACTGCACTCTGAGAGAGCACTTGTG 540  
Db 560 GAGGAGATTAAGTGGCGCACTCCCGGTGACATCACTGCACTCTGAGAGAGCACTTGTG 619  
QY 541 AAGTCAGAGCTGCTACACATTTGATTCAGAGTCTTGAAGATTCGATGATTCAGGA 600  
Db 620 AAGTCAGAGCTGCTACACATTTGATTCAGAGTCTTGAAGATTCGATGATTCAGGA 679  
QY 601 AGAAGGATTTCTGAGTTTGAAGCAAAATGCAATGCAAGCAAGCAAGCAAGCAAGTGTG 660  
Db 680 AGAAGGATTTCTGAGTTTGAAGCAAAATGCAATGCAAGCAAGCAAGCAAGTGTG 739  
QY 661 ATTGAAGCTTTGAGTCTCACTTACCCGTTCCGTTACCCCACTCTCAAGAGCTCTCTG 720  
Db 740 ATTGAAGCTTTGAGTCTCACTTACCCGTTCCGTTACCCCACTCTCAAGAGCTCTCTG 799  
QY 721 AATTTCTGATTTACCCCAAGATTTTACAGGTATCTGACAGAGTCTCTTGGCCAG 780  
Db 800 AATTTCTGATTTACCCCAAGATTTTACAGGTATCTGACAGAGTCTCTTGGCCAG 859  
QY 781 GAGGAGGCAAGATTTCTGATCTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 840  
Db 860 GAGGAGGCAAGATTTCTGATCTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 919  
QY 841 GCAATTTCACTTACTAGATGATGATGATGATGATGATGATGATGATGATGATGATG 900  
Db 920 GCAATTTCACTTACTAGATGATGATGATGATGATGATGATGATGATGATGATGATG 979  
QY 901 TCAAAATCAAGATTTTCTTATCAGGCTGAGATGCTTCAAGGCTGATCTGCTTAAAG 960  
Db 980 TCAAAATCAAGATTTTCTTATCAGGCTGAGATGCTTCAAGGCTGATCTGCTTAAAG 1039  
QY 961 GATTTGAGGATTAAGGCTTATCCTCAAGAGCTGAGATGCTTCAAGGCTGATCTGCTTAAAG 1020  
Db 1040 GATTTGAGGATTAAGGCTTATCCTCAAGAGCTGAGATGCTTCAAGGCTGATCTGCTTAAAG 1099  
QY 1021 GTCCTTTGAAAATTAAGGAGACACAAGAAAGAAAGAGCTTACCCAGCATATA 1080  
Db 1100 GTCCTTTGAAAATTAAGGAGACACAAGAAAGAAAGAGCTTACCCAGCATATA 1159  
QY 1081 CCGGCGGAGATTTCTCTCAAGTCAATTTTAACTGCTGCTTGAAGATTCGAGCAATTC 1140  
Db 1160 CCGGCGGAGATTTCTCTCAAGTCAATTTTAACTGCTGCTTGAAGATTCGAGCAATTC 1219  
QY 1141 AAAAAGGCAATTTTGGAGGCGCTTGTGACATTAACGATGACATGCTGAAAAGCCGAG 1200  
Db 1220 AAAAAGGCAATTTTGGAGGCGCTTGTGACATTAACGATGACATGCTGAAAAGCCGAG 1279  
QY 1201 CTACAGAGCTGTCAGATTAACAAGGAGGAGCCGATATATAGCCGCTTGTACGAGATGCC 1260  
Db 1280 CTACAGAGCTGTCAGATTAACAAGGAGGAGCCGATATATAGCCGCTTGTACGAGATGCC 1339  
QY 1261 TGGGCTGCTTTGATGATCTCTCTGCTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCT 1320  
Db 1340 TGGGCTGCTTTGATGATCTCTCTGCTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCT 1399  
QY 1321 CTGCTGGAATCTTCTTAACTTCAACCAAGACCAATTCGCTGCAACTCAAGTTTA 1380  
|||||

Db 1400 CTGCTGGAATCTTCTTAACTTCAACCAAGACCAATTCGCTGCAACTCAAGTTTA 1459  
QY 1381 TTTCAACCCAGGAAGAGCTCCATTTTGTCTTCAACATTTGTGGAATTTGTCTCACTGCCA 1440  
Db 1460 TTTCAACCCAGGAAGAGCTCCATTTTGTCTTCAACATTTGTGGAATTTGTCTCACTGCCA 1519  
QY 1441 ACAAGGTTCTGCGGAAGGAGATATGATACAGGCTGCGCTTGTGATTTGTCTTCAAGTT 1500  
Db 1520 ACAAGGTTCTGCGGAAGGAGATATGATACAGGCTGCGCTTGTGATTTGTCTTCAAGTT 1579  
QY 1501 CTTTCAAGCAAACTATCATGATATCCCATGAAAGACAGGCGGAAAGCCCTGCTCTTAAGATA 1560  
Db 1580 CTTTCAAGCAAACTATCATGATATCCCATGAAAGACAGGCGGAAAGCCCTGCTCTTAAGATA 1639  
QY 1561 TCCATCTCTCTGGAACCAAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
Db 1640 TCCATCTCTCTGGAACCAAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1699  
QY 1621 ATAATGTTGGTTCAGGAACCGGCAATAGCCCGTTTATTTGGTTCTTCAACAATAGAG 1680  
Db 1700 ATAATGTTGGTTCAGGAACCGGCAATAGCCCGTTTATTTGGTTCTTCAACAATAGAG 1759  
QY 1681 AAATCTCAAGAAACAACCAAGATGAAATTTTGGAGCAATG---GTTTTTGGCTGC 1737  
Db 1760 AAATCTCAAGAAACAACCAAGATGAAATTTTGGAGCAATGTTGTTTGGTCTGC 1819  
QY 1738 AGGCAATTAAGATAGGATTAATCTATTCAAGAAAGAGCTCAGACATTTCTTAAAGATGG 1797  
Db 1820 AGGCAATTAAGATAGGATTAATCTATTCAAGAAAGAGCTCAGACATTTCTTAAAGATGG 1879  
QY 1798 ATCTTAATCTATTAAGGTTTCTTCTCAAGATGCTCTGTTGGGAGAGAGAGGCC 1857  
Db 1880 ATCTTAATCTATTAAGGTTTCTTCTCAAGATGCTCTGTTGGGAGAGAGAGGCC 1939  
QY 1858 CCAAGCAAGATTAATTAACAACAATCCAGCTTCAAGGCGGAGAGAGAGATCTC 1917  
Db 1940 CCAAGCAAGATTAATTAACAACAATCCAGCTTCAAGGCGGAGAGAGAGATCTC 1999  
QY 1918 CTCACAGAGAAAGGCAATTTATTTGTGTGAGATGCAAAAGATTAAGCAAGATGTA 1977  
Db 2000 CTCACAGAGAAAGGCAATTTATTTGTGTGAGATGCAAAAGATTAAGCAAGATGTA 2059  
QY 1978 CATGATGCTCTTGTGCAAAATTAATTAAGCAAGAGGTTGAGTTGAAAACTTAAGCAATG 2037  
Db 2060 CATGATGCTCTTGTGCAAAATTAATTAAGCAAGAGGTTGAGTTGAAAACTTAAGCAATG 2119  
QY 2038 AAAACCTGCGCACTTTAAAAGAAAGAAAGCTTACCTTCAAGATATTTGCTATA 2094  
Db 2120 AAAACCTGCGCACTTTAAAAGAAAGAAAGCTTACCTTCAAGATATTTGCTATA 2176

RESULT 2  
AX050463 3259 bp DNA linear PAT 12-JAN-2001

LOCUS AX050463  
DEFINITION Sequence 23 from Patent WO0071754.

ACCESSION AX050463  
VERSION AX050463.1 GI:12226668

KEYWORDS  
SOURCE  
ORGANISM

Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

#### REFERENCE

1 Johnson, W.G. and Stenroos, E.S.  
Methods for diagnosing, preventing, and treating developmental  
disorders due to a combination of genetic and environmental factors  
Patent: WO 0071754-A 23 NOV-2000;

JOURNAL  
University of Medicine and Dentistry of New Jersey (US)  
Location/Qualifiers

#### FEATURES

1..3259  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"



## ORIGIN

Query Match 88.5%; Score 1854; DB 6; Length 3259;  
Beet Local Similarity 99.9%; Pred. No. 0;  
Matches 2094; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

1 ATGAGAGAGTTCTGTACTATATGCTACACAGCAGGAGCAGGCAAAAGCCATCCGAGAA 60  
80 ATGAGAGAGTTCTGTACTATATGCTACACAGCAGGAGCAGGCAAAAGCCATCCGAGAA 139  
61 GAAATGTGTAGCAAGCTGTGTGATGATGATTTCTGCGATCTTCACTGTATTTAGTGA 120  
140 GAAATGTGTAGCAAGCTGTGTGATGATGATTTCTGCGATCTTCACTGTATTTAGTGA 199  
121 TCCGATAGTATGACTTAAACCGGAAAGCTCCCTGTGTGTGTGTCTTCAACG 180  
200 TCCGATAGTATGACTTAAACCGGAAAGCTCCCTGTGTGTGTGTCTTCAACG 259  
181 GGCACCGGAGACCCACCCGACACAGCCGCAAGTTGTTAGGAAATACAGAACCAACA 240  
260 GGCACCGGAGACCCACCCGACACAGCCGCAAGTTGTTAGGAAATACAGAACCAACA 319  
241 CTGCCGTTGATTTCTTCTCTCACTGCGGTATGAGTTCTGCGTCTCGTATTCAGAA 300  
320 CTGCCGTTGATTTCTTCTCTCACTGCGGTATGAGTTCTGCGTCTCGTATTCAGAA 379  
301 TACACCTTACTTTTGCATATGGGGGGAAGATTTGATTAACGACTTCAAGAGCTTGAAGC 360  
380 TACACCTTACTTTTGCATATGGGGGGAAGATTTGATTAACGACTTCAAGAGCTTGAAGC 439  
361 CGGCAATTTCTATGACACTGACATGACATGACTGTGTGTGTGTGTGTGTGTGTGTGT 420  
440 CGGCAATTTCTATGACACTGACATGACATGACTGTGTGTGTGTGTGTGTGTGTGTGT 499  
421 CGTGGATTTGCTGACTCTGCGCAAGCCCTCAGAAAGCATTTTATGCTCAAGCAGAGCAA 480  
500 CGTGGATTTGCTGACTCTGCGCAAGCCCTCAGAAAGCATTTTATGCTCAAGCAGAGCAA 559  
481 GAGGAGATTAAGTGGGCACTCCCGGTGGATCACTCTGCACTCTTGAAGGACAGACCTTGG 540  
560 GAGGAGATTAAGTGGGCACTCCCGGTGGATCACTCTGCACTCTTGAAGGACAGACCTTGG 619  
541 AAGTCAGAGCTCTACATTTGAATCTCAAGTCAAGCTCTGAGATTCATGATTTCAAGA 600  
620 AAGTCAGAGCTCTACATTTGAATCTCAAGTCAAGCTCTGAGATTCATGATTTCAAGA 679  
601 AGAAAGATTTCTGAGGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATCCAATGTTGA 660  
680 AGAAAGATTTCTGAGGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATCCAATGTTGA 739  
661 ATTGAAGACTTTGAGTCTCTCACTTACCGGTGGATACCCCACTCTCAAGCTTCTG 720  
740 ATTGAAGACTTTGAGTCTCTCACTTACCGGTGGATACCCCACTCTCAAGCTTCTG 799  
721 AATATTCCTGTTTACCCCAAGATTTTACAGGTCATCTGAGGAGTCTTGGCCAG 780  
800 AATATTCCTGTTTACCCCAAGATTTTACAGGTCATCTGAGGAGTCTTGGCCAG 859  
781 GAGGAAAGCAAGTATCTGTGACTTCAAGAGATCCAGTTTTCAGTGCCTTCAAG 840  
860 GAGGAAAGCAAGTATCTGTGACTTCAAGAGATCCAGTTTTCAGTGCCTTCAAG 919  
841 GCAATTCATTTCTGCAATGATGCTCAATTAACCACTCTGCTGTGTGATTTGAAGATT 900  
920 GCAATTCATTTCTGCAATGATGCTCAATTAACCACTCTGCTGTGTGATTTGAAGATT 979  
901 TCAATATACAGACTTTTCTATGAGCTGTGAGATGCTTCAAGGATTCGCTTCAAG 960  
980 TCAATATACAGACTTTTCTATGAGCTGTGAGATGCTTCAAGGATTCGCTTCAAG 1039  
961 GATTTCTGAGGATCAAAAGCTTCTCAAGATCTGCAAGCTTGAAGATTAAGAGAGCACTGC 1020  
1040 GATTTCTGAGGATCAAAAGCTTCTCAAGATCTGCAAGCTTGAAGATTAAGAGAGCACTGC 1099

1021 GTCTTTTGAATAAAGCAGACAAAGAAAGAGCTACCTTACCCGACATATA 1080  
1100 GTCTTTTGAATAAAGCAGACAAAGAAAGAGCTACCTTACCCGACATATA 1159  
1081 CCGCGGAGATGTTCTCTCAAGTTCAATTTTACCTGTGTCTTGAATCCGACATTTCT 1140  
1160 CCGCGGAGATGTTCTCTCAAGTTCAATTTTACCTGTGTCTTGAATCCGACATTTCT 1219  
1141 AAAAAAGATTTTTCGAGCCCTGTGAGCTATACAGTGAACAGTGTGAAGAGCAGG 1200  
1220 AAAAAAGATTTTTCGAGCCCTGTGAGCTATACAGTGAACAGTGTGAAGAGCAGG 1279  
1201 CTACAGAGCTGTGAGTAAACAAGGGCAGCCGATTAAGCCGTTTGTACAGATGCC 1260  
1280 CTACAGAGCTGTGAGTAAACAAGGGCAGCCGATTAAGCCGTTTGTACAGATGCC 1339  
1261 TGTGCTCTGTGTGATCTCTCTCGCTTCCCTTCTGTCAGCCACACTCAGTCTC 1320  
1340 TGTGCTCTGTGTGATCTCTCTCGCTTCCCTTCTGTCAGCCACACTCAGTCTC 1399  
1321 CTGCTCGAATCTTCTCTTAACTTCAACCAAGCAATTAAGTGTGCAAGCTCAAGTTA 1380  
1400 CTGCTCGAATCTTCTCTTAACTTCAACCAAGCAATTAAGTGTGCAAGCTCAAGTTA 1459  
1381 TTTGACCCAGGAAGCTCAATTTGTCTTCAACATTTGTGAATTTCTGTCTACTGCCACA 1440  
1460 TTTGACCCAGGAAGCTCAATTTGTCTTCAACATTTGTGAATTTCTGTCTACTGCCACA 1519  
1441 ACAGAGTTCTGCGAGAGGATATGACAGGCTGCGCTGCTGTGTGTGCTTCAATT 1500  
1520 ACAGAGTTCTGCGAGAGGATATGACAGGCTGCGCTGCTGTGTGTGCTTCAATT 1579  
1501 CTTGAGCCAAATATATGATATCCATGAAGACAGCGGAAAGCCCTGCTCTTAAGATA 1560  
1580 CTTGAGCCAAATATATGATATCCATGAAGACAGCGGAAAGCCCTGCTCTTAAGATA 1639  
1561 TCCATCTCTCTGGAACAAATTTCTTCCACTTACAGATGACCCCTCAATCCCATC 1620  
1640 TCCATCTCTCTGGAACAAATTTCTTCCACTTACAGATGACCCCTCAATCCCATC 1699  
1621 ATATATGTGTGCTCAAGAACCGGCAATAGCCCGTTATTTGAGTTCTTCAACATTAAGAG 1680  
1700 ATATATGTGTGCTCAAGAACCGGCAATAGCCCGTTATTTGAGTTCTTCAACATTAAGAG 1759  
1681 AAATCTCAAGAACCAACCCAGATGGAATTTTGAAGCAATGTG--GTTTTTGGCTGC 1737  
1760 AAATCTCAAGAACCAACCCAGATGGAATTTTGAAGCAATGTGTTTGTGCTGC 1819  
1738 AGGCAATTAAGGATTAATCTATTCAGAAAGAGCTCAGACATTTCTTAAAGCATGGG 1797  
1820 AGGCAATTAAGGATTAATCTATTCAGAAAGAGCTCAGACATTTCTTAAAGCATGGG 1879  
1798 ATCTTAATCATCTTAAGGTTCTTCTCAAGAGATGCTCTGTGTGGAGAGAGAGCC 1857  
1880 ATCTTAATCATCTTAAGGTTCTTCTCAAGAGATGCTCTGTGTGGAGAGAGAGCC 1939  
1858 CCAAGAAAGTATGTAACAAGCAATCCAGCTTCAATGTCAGAGGTGCGGAGATCTTC 1917  
1940 CCAAGAAAGTATGTAACAAGCAATCCAGCTTCAATGTCAGAGGTGCGGAGATCTTC 1999  
1918 CTCCAGGAGAAAGGCATATTTATGTTGTGAGATGCAAGAAATATGAGCAAGATGTA 1977  
2000 CTCCAGGAGAAAGGCATATTTATGTTGTGAGATGCAAGAAATATGAGCAAGATGTA 2059  
1978 CATGATGCCCTTGTCAATTAATTAAGCAAGAGGTTGAGGTTGAAGAACTAAGCAATG 2037  
2060 CATGATGCCCTTGTCAATTAATTAAGCAAGAGGTTGAGGTTGAAGAACTAAGCAATG 2119  
2038 AAAACCTTGGCCACTTTTAAAGAAAGAAAGCCCTTCAAGATATTTGTGCTATTA 2094  
2120 AAAACCTTGGCCACTTTTAAAGAAAGAAAGCCCTTCAAGATATTTGTGCTATTA 2176

RESULT 3  
AF025794 3259 bp mRNA linear PRI 26-MAR-1998  
LOCUS Homo sapiens methionine synthase reductase (MTRR) mRNA, complete  
DEFINITION  
ACCESSION AF025794  
VERSION AF025794.1 GI:2981302  
SOURCE Homo sapiens (human)  
KEYWORDS Bukavota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 3259)  
AUTHORS Leclerc, D., Wilson, A., Dumas, R., Gafuick, C., Song, D., Watkins, D., Heng, H.H.Q., Rommens, J.M., Scherer, S.W., Rosenblatt, D.S. and Gravel, R.A.  
TITLE Cloning and mapping of a cDNA for methionine synthase reductase, a flavoprotein defective in patients with homocystinuria  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (6), 3059-3064 (1998)  
MEDLINE 98169496  
PUBMED 9501215  
REFERENCE 2 (bases 1 to 3259)  
AUTHORS Leclerc, D.  
TITLE Direct Submission  
JOURNAL Submitted (19-SEP-1997) Human Genetics, McGill University - Montreal Children's Hospital Research Institute, 4060 Ste-Catherine West, Montreal, Que H3Z 2Z3, Canada  
REFERENCE 3 (bases 1 to 3259)  
AUTHORS Leclerc, D.  
TITLE Direct Submission  
JOURNAL Submitted (12-NOV-1997) Human Genetics, McGill University - Montreal Children's Hospital Research Institute, 4060 Ste-Catherine West, Montreal, Que H3Z 2Z3, Canada  
REMARK Sequence update by submitter  
FEATURES  
source location/Qualifiers  
1..3259  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/map="5p15.2-p15.3"  
1..3259  
/gene="MTRR"  
80..2176  
/gene="MTRR"  
/note="B12 vitamin; cob complementation group"  
/codon\_start=1  
/product="methionine synthase reductase"  
/protein\_id="AAC39667.1"  
/db\_xref="GI:2981303"  
/translation="MRRFLLYVATQOGAKAIABEMCEQAVHGFSAADLHCSISDXY  
DIKETAPLVVVVSTGTGDPDPARKVEIONOTLPVDFPHLRGVLGIDSEYK  
YFNGGALPVPASPLRTDLYKSELHIESQVELRPDSDGRKDSVTKQNAVNSGN  
VAVLEDFESSLTRSVPLSQSLNPGLPPEYLVQHLQESLGQESQVSVTSADPVQY  
PIBAVOLITNDALKITLVLELDISNTDFSYQPDASFVCPNSDSVOSILQRLQIE  
DKBEHCVLTKIKADTKKGAATLPOHIFAGSIQIFPWCLEAIPKAPALAVQY  
SDAEKRYLQELKOGADSVRFVADACGLDILAPSCOPPISTLEHLPKIQP  
RYSQASSSLPHQKLAHFNIVETPSTATTEVLRKQVCTGCMALIVASTQNYHS  
HEBSGRALAKRISISPTTNSFHLPPDPSPILIVGCGTGLAPFGLQREKIQEON  
VDONIQHGOQVFAIRLQENDNYLFRKELRHLKKGILTHLVFSRDAVVEEBEAPAY  
LATLEKEKRYLQIDIMS"

ORIGIN  
Query Match 88.5%; Score 1854; DB 9; Length 3259;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2094; Conservative 0; Mismatches 0; Indels 3; Gaps 1;  
QY 1 ATGAGAGGTTCTGTTACTATGCTACACAGAGGAGACAGAAAGGCATGCGAGAA 60  
Db 80 ATGAGAGGTTCTGTTACTATGCTACACAGAGGAGACAGAAAGGCATGCGAGAA 139

QY 61 GAAATGTGAGCAAGCTGTGTGATGATTTTTCGACGATCTTCACTGTATTAAGTGA 120  
Db 140 GAAATGTGAGCAAGCTGTGTGATGATTTTTCGACGATCTTCACTGTATTAAGTGA 199  
QY 121 TCCGATTAAGTATGACCTTAACAAACCGGAAACAGCCCTCTGTTGTTGTGTTCTACAG 180  
Db 200 TCCGATTAAGTATGACCTTAACAAACCGGAAACAGCCCTCTGTTGTTGTGTTCTACAG 259  
QY 181 GGCACCGAGACCAACCCGACACAGCCCGCAAGTTGTTAAGAAATACAGAACCAACA 240  
Db 260 GGCACCGAGACCAACCCGACACAGCCCGCAAGTTGTTAAGAAATACAGAACCAACA 319  
QY 241 CTGCGGTTGATTTCTTTGCTCACTGCGGTATGGGTTACTGGGTCTGGTATTCAGA 300  
Db 320 CTGCGGTTGATTTCTTTGCTCACTGCGGTATGGGTTACTGGGTCTGGTATTCAGA 379  
QY 301 TACACCTACTTTTGGCAATGGGGGGAAGATTAATTGAATAACGCTTCAAGAGCTGAGGC 360  
Db 380 TACACCTACTTTTGGCAATGGGGGGAAGATTAATTGAATAACGCTTCAAGAGCTGAGGC 439  
QY 361 CGGCAATTTCTATGACACTGACATGACATGACATGACATGACATGACATGACATGACAT 420  
Db 440 CGGCAATTTCTATGACACTGACATGACATGACATGACATGACATGACATGACATGACAT 499  
QY 421 CCGTGATTTCTGACTCTGCGCCAGCCCTCAGAAAGCATTTTATGTCAGAGAGACAA 480  
Db 500 CCGTGATTTCTGACTCTGCGCCAGCCCTCAGAAAGCATTTTATGTCAGAGAGACAA 559  
QY 481 GAGGATATAGGGGCGCACTCCCGGNGCATACCGTCGATCCCTGAGGACAGACCTGTG 540  
Db 560 GAGGATATAGGGGCGCACTCCCGGNGCATACCGTCGATCCCTGAGGACAGACCTGTG 619  
QY 541 AAGTCAGACTGCTACACATTTGAAAGCAAAATGCAAGTGAACGCAACCAATTCATTTGTA 600  
Db 620 AAGTCAGACTGCTACACATTTGAAAGCAAAATGCAAGTGAACGCAACCAATTCATTTGTA 679  
QY 601 AGAAAGATTTGAGGTTTGAAGCAAAATGCAAGTGAACGCAACCAATTCATTTGTA 660  
Db 680 AGAAAGATTTGAGGTTTGAAGCAAAATGCAAGTGAACGCAACCAATTCATTTGTA 739  
QY 661 ATGGAAGCTTGAGTCCCTCACTTACCGTGGGTACCCCACTCCACAAGCTCTG 720  
Db 740 ATGGAAGCTTGAGTCCCTCACTTACCGTGGGTACCCCACTCCACAAGCTCTG 799  
QY 721 AATATTCCTGTTTACCCCAAGATATTTACAGTACATCTGACGAGTCTCTGGCCAG 780  
Db 800 AATATTCCTGTTTACCCCAAGATATTTACAGTACATCTGACGAGTCTCTGGCCAG 859  
QY 781 GAGGAAAGCCAGATCTGTGACTTCAGCAGATCCAGTTTCAAGTCCCAATTTCAAG 840  
Db 860 GAGGAAAGCCAGATCTGTGACTTCAGCAGATCCAGTTTCAAGTCCCAATTTCAAG 919  
QY 841 GCGATTCATCTACGAAATGATGCGCATTAACCACTGCTGTGATGAATTTGACAT 900  
Db 920 GCGATTCATCTACGAAATGATGCGCATTAACCACTGCTGTGATGAATTTGACAT 979  
QY 901 TCAATAACAGACTTTCTATCAGCTGAGATGCTTCAAGCTGATCTGCTTACAGT 960  
Db 980 TCAATAACAGACTTTCTATCAGCTGAGATGCTTCAAGCTGATCTGCTTACAGT 1039  
QY 961 GATTTGAGTACAAAGCTTCTCAAGAGCTGACCTTGAAGATTAAGAGAGCACTGC 1020  
Db 1040 GATTTGAGTACAAAGCTTCTCAAGAGCTGACCTTGAAGATTAAGAGAGCACTGC 1099  
QY 1021 GTGCTTTGAAATTAAGGAGACAAAGAAAGAGAGCTTACCTTACCCGACATATA 1080  
Db 1100 GTGCTTTGAAATTAAGGAGACAAAGAAAGAGAGCTTACCTTACCCGACATATA 1159  
QY 1081 CTGCGGAGATTTCTCTCAGATTCATTTTATCTGTGTTGTAATCCGAGCAATTTCT 1140  
Db 1160 CTGCGGAGATTTCTCTCAGATTCATTTTATCTGTGTTGTAATCCGAGCAATTTCT 1219

1141 AAAAGGCAATTTTGGAGCCCTTGTGACATATACAGTGAAGTGTGAAAAAGCGACG 1200  
1220 AAAAGGCAATTTTGGAGCCCTTGTGACATATACAGTGAAGTGTGAAAAAGCGACG 1279  
1201 CTACAGAGCTGTGACGTAACAAAGGGGACCGGATTAAGCCGTTTGTACAGATGCGC 1260  
1280 CTACAGAGCTGTGACGTAACAAAGGGGACCGGATTAAGCCGTTTGTACAGATGCGC 1339  
1261 TGTGCTGTGTGTGATCTCTCCCGCTTCCCTTCTTGGCAAGCAGACCTAGCTTC 1320  
1340 TGTGCTGTGTGTGATCTCTCCCGCTTCCCTTCTTGGCAAGCAGACCTAGCTTC 1399  
1321 CTGCTGGAACATCTTCTTAACTTCAACCCAGACATATTTGCTGACAGCTCAAGTTA 1380  
1400 CTGCTGGAACATCTTCTTAACTTCAACCCAGACATATTTGCTGACAGCTCAAGTTA 1459  
1381 TTTTCAACCGAAGAAAGCTCCATTTTGTCTTCAACATTTGTGAAATTTCTGTCTAGTCCACA 1440  
1460 TTTTCAACCGAAGAAAGCTCCATTTTGTCTTCAACATTTGTGAAATTTCTGTCTAGTCCACA 1519  
1441 ACAAGAGTTCTGGGAGGAGATATGTAAGGCTGGCTGCTGTTGTTGCTTCAAGTT 1500  
1520 ACAAGAGTTCTGGGAGGAGATATGTAAGGCTGGCTGCTGTTGTTGCTTCAAGTT 1579  
1501 CTTCAGCCAAATATATGATGATCCCATGAAAGACAGGGGAAAGCCCTGGCTCTAAGATA 1560  
1580 CTTCAGCCAAATATATGATGATCCCATGAAAGACAGGGGAAAGCCCTGGCTCTAAGATA 1639  
1561 TCCATCTCTCTGACAAACAATTTCTTCACTTCAAGATGACCCCTCAATCCCACATC 1620  
1640 TCCATCTCTCTGACAAACAATTTCTTCACTTCAAGATGACCCCTCAATCCCACATC 1699  
1621 ATATAGTGGGTCACAGAAACCGGACATAGCCGCTTATTTGGGTTCTTCAACAATAGAG 1680  
1700 ATATAGTGGGTCACAGAAACCGGACATAGCCGCTTATTTGGGTTCTTCAACAATAGAG 1759  
1681 AAATCTCAAGAAACAACCCAGATGGAATTTTGGAGCAATGNG---GTTTTTGGCTGC 1737  
1760 AAATCTCAAGAAACAACCCAGATGGAATTTTGGAGCAATGNGTTTTTGGCTGC 1819  
1738 AGGATTAAGGATGAGATTAATCTATTCAGAAAAGAGCTCAGACATTTCTTAAAGATGGG 1797  
1820 AGGATTAAGGATGAGATTAATCTATTCAGAAAAGAGCTCAGACATTTCTTAAAGATGGG 1879  
1798 ATCTTAATCATCTTAAAGGTTTCTTCTCAAGAGATGCTCTGTTGGGGAAGAGAGCC 1857  
1880 ATCTTAATCATCTTAAAGGTTTCTTCTCAAGAGATGCTCTGTTGGGGAAGAGAGCC 1939  
1858 CCAAGAAAGTATGTAACAACAATTCAGACTTCATGCGCAGAGGTGGCGAGATCTTC 1917  
1940 CCAAGAAAGTATGTAACAACAATTCAGACTTCATGCGCAGAGGTGGCGAGATCTTC 1999  
1918 CTCACAGAGAAAGGCAATATTTATGTTGTGAGATGCAAAAGATATGCGCAAGATGTA 1977  
2000 CTCACAGAGAAAGGCAATATTTATGTTGTGAGATGCAAAAGATATGCGCAAGATGTA 2059  
1978 CATGATGCCCTTGTGCAATATATAGCAAAAGAGTTGAGTTGAAAACTAGAAGCATG 2037  
2060 CATGATGCCCTTGTGCAATATATAGCAAAAGAGTTGAGTTGAAAACTAGAAGCATG 2119  
2038 AAAACCTGGCCACTTTAAAAAGAAAAAGCGTACTCTTCAAGATATTTGGTCAATA 2094  
2120 AAAACCTGGCCACTTTAAAAAGAAAAAGCGTACTCTTCAAGATATTTGGTCAATA 2176

RESULT 4  
AF121214 3291 bp mRNA linear PRI 12-DEC-1999  
LOCUS AF121214 Homo sapiens methionine synthase reductase (MTRR) mRNA, complete  
DEFINITION cds.  
ACCESSION AF121214  
VERSION AF121214.1 GI:6561338  
KEYWORDS

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS  
1 (bases 1 to 3291)  
Lecierc,D., Odleyre,M., Wu,Q., Wilson,A., Huizenga,J.J., Rozen,R.,  
Scherer,S.W. and Gravel,R.A.  
TITLE  
Molecular cloning, expression and physical mapping of the human  
methionine synthase reductase gene  
JOURNAL  
Gene 240 (1), 75-88 (1999)  
MEDLINE  
20033550  
PUBMED  
10564814  
REFERENCE  
2 (bases 1 to 3291)  
Lecierc,D., Odleyre,M.-H., Wu,Q., Wilson,A., Huizenga,J.J.,  
Johns,T., Shoubiridge,E.A., Rosenblatt,D.S., Scherer,S.W., Rozen,R.  
and Gravel,R.A.  
TITLE  
Direct Submision  
JOURNAL  
Submitted (18-JAN-1999) Human Genetics, Montreal Children's  
Hospital, 4060 Ste-Catherine West, Montreal, Quebec H3Z 2Z3, Canada  
FEATURES  
source  
1..3291  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/map="5p15.2-p15.3; between W11755 and D5S1957"  
1..3291  
/gene="MTRR"  
1..30  
/gene="MTRR"  
/evidence=not\_experimental  
31..3208  
/gene="MTRR"  
/EC\_number="2.1.1.135"  
/function="maintenance of methionine synthase in an active  
state"  
/note="alternatively spliced; alternative splicing results  
in extended 5' sequence compared to GenBank Accession  
Number AF025794; B12 vitamin; cblE complementation group"  
/codon\_start=1  
/product="methionine synthase reductase"  
/protein\_id="AA16876.1"  
/db\_xref="GI:6561339"  
/translation="MGASVPRAGALVVALVCSPTVTCLEVRERFLIYATGQOAKA  
IAEMCEQAVHGFADLHICISBSKTYLKTETAPLVVYVSTGTGDPDTRARKYKE  
IONQTLVPDFALHLYGILGDSRYTFCEGKILDRLOELGALHFDYDGHADCV  
GLELVBPWINGLWALRKHFSSRSGEISGALPVASPLRTDLVNSBLHISQV  
ELLRPDGSRKDSVTLKONAVNSQNSVVIDPESLSTRSVPLSQASLIDISLTPRY  
LOVHIOESLGOEESQVSTVSDPVPOVLSRAVOLTTNDALIKTLLVELNDISLTPRY  
OQGDAPVTCPSNDSRVSQLOLQLEDRERHCYLKIKADTRKKGATLPQHPAGCS  
LQPIFTWCELEIATPKAPLRLVDYTSDAKRLQELCSKQADYSRFVDRAC  
LIDLILAPSPCCPPLSLLEHLPKLQPPRYSCASSLPHPGLHVFNLVELESTATT  
EVLARKGVCQGMALILVASVLQPINIASHSDSGKALAPKISLSPRTNSPHLDPSIP  
ILWVGEGTIAPIFGLQREKLOHDPNGSGAMLPFGCHKRDYLFRELEHFL  
KGGITLHKVSPSRDAPVGEERAARKYVODNTQLHGOVARYLLDENGHIYVCGDAKN  
NAKVDHDLVQILISKEVGEKLEAKTLATLKEBRYLDIYS"

transit\_peptide  
31..123  
/gene="MTRR"  
/evidence=not\_experimental  
3' UTR  
2209..3291  
/gene="MTRR"  
/evidence=not\_experimental  
polyA\_signal  
3246..3251  
/gene="MTRR"  
/evidence=not\_experimental  
polyA\_site  
3268  
/gene="MTRR"  
/evidence=experimental  
polyA\_site  
3276  
/gene="MTRR"  
/evidence=experimental  
ORIGIN

Query Match 88.5%; Score 1854; DB 9; Length 3291;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2094; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 ATGAGAGGTTTCTGTTACTATATGCTACAGAGAGGACAGGCAAGCCATCGAGAA 60  
DB 112 ATGAGAGGTTTCTGTTACTATATGCTACAGAGAGGACAGGCAAGCCATCGAGAA 171  
QY 61 GAAATGTGAGCAAGCTGTGTATCATGATTTTCTGAGATCTTCACTGTATTAGTAA 120  
DB 172 GAAATGTGAGCAAGCTGTGTATCATGATTTTCTGAGATCTTCACTGTATTAGTAA 231  
QY 121 TCGGATAGATGACCTTAACCAACGAAACAGCTCCCTGTGTGTGTGTCTTACACG 180  
DB 232 TCGGATAGATGACCTTAACCAACGAAACAGCTCCCTGTGTGTGTGTCTTACACG 291  
QY 181 GGCACCGAGACCCACCCGACAGAGCCGCAAGTTTGTAAAGAAATACAGAACCAACA 240  
DB 292 GGCACCGAGACCCACCCGACAGAGCCGCAAGTTTGTAAAGAAATACAGAACCAACA 351  
QY 241 CTGCGCGTGTATTTCTTTGCTCACTGCGGTATGGGTTACTGGGCTCGGTATTCAGAA 300  
DB 352 CTGCGCGTGTATTTCTTTGCTCACTGCGGTATGGGTTACTGGGCTCGGTATTCAGAA 411  
QY 301 TACACCTACTTTTGCAATGGGGGGAAGATATGATTAAGACCTTCAAGGCTTGAAGCC 360  
DB 412 TACACCTACTTTTGCAATGGGGGGAAGATATGATTAAGACCTTCAAGGCTTGAAGCC 471  
QY 361 CGGCAATTTCTATGACACTGACATGACATGATGACTGTATAGTTTGAACCTTGTGTAG 420  
DB 472 CGGCAATTTCTATGACACTGACATGACATGATGACTGTATAGTTTGAACCTTGTGTAG 531  
QY 421 CCGTGATTTCTGACCTCTGCGCAAGCCCTCAAGAAAGATTTTAAAGTCAACAGAGACAA 480  
DB 532 CCGTGATTTCTGACCTCTGCGCAAGCCCTCAAGAAAGATTTTAAAGTCAACAGAGACAA 591  
QY 481 GAGAGATAGTGGCGGACCTCCGCTGGGACATCACTGCACTCTTGAAGAGACCTTGTG 540  
DB 592 GAGAGATAGTGGCGGACCTCCGCTGGGACATCACTGCACTCTTGAAGAGACCTTGTG 651  
QY 541 AAGTCAGAGCTGCTACATCAATGATCTCAAGTGAAGCTTCTGAGATTCATGATTCAGA 600  
DB 652 AAGTCAGAGCTGCTACATCAATGATCTCAAGTGAAGCTTCTGAGATTCATGATTCAGA 711  
QY 601 AGAAGGATTTCTGAGGTTTGAAGCAAAATGACGTGAACAGCAACCAATCCAATGTTGA 660  
DB 712 AGAAGGATTTCTGAGGTTTGAAGCAAAATGACGTGAACAGCAACCAATCCAATGTTGA 771  
QY 661 ATTGAAGCTTTGATGCTCTCACTTACCCGTTGGGTACCCCACTCTCAAGGCTCTCTG 720  
DB 772 ATTGAAGCTTTGATGCTCTCACTTACCCGTTGGGTACCCCACTCTCAAGGCTCTCTG 831  
QY 721 AATATTCCTGGTTTACCCCGAATATTTACAGGTATCATCTGACAGAGTCTCTTGCGCAG 780  
DB 832 AATATTCCTGGTTTACCCCGAATATTTACAGGTATCATCTGACAGAGTCTCTTGCGCAG 891  
QY 781 GAGAGAAAGCAAGTATCTGACTTTCAGAGATTCAGGTTTTCAGAGTCCCAATTTCAAG 840  
DB 892 GAGAGAAAGCAAGTATCTGACTTTCAGAGATTCAGGTTTTCAGAGTCCCAATTTCAAG 951  
QY 841 GCGATTTCAATTTACTAGGAATGATGCAATAAAACCACTGCTGCTGATGAATTTGACATT 900  
DB 952 GCGATTTCAATTTACTAGGAATGATGCAATAAAACCACTGCTGCTGATGAATTTGACATT 1011  
QY 901 TCAAAATCAGACTTTTCTTATAGCTCTGAGATCTTCAAGCGTATCTGCTTAAAGT 960  
DB 1012 TCAAAATCAGACTTTTCTTATAGCTCTGAGATCTTCAAGCGTATCTGCTTAAAGT 1071  
QY 961 GATTTGAGGTACAAAGCTTATCTCAAAAGCTGACGTTGAAGTAATAAGAGACACTGC 1020  
DB 1072 GATTTGAGGTACAAAGCTTATCTCAAAAGCTGACGTTGAAGTAATAAGAGACACTGC 1131  
QY 1021 GTCTTTTGAATAAAGGACACAAAGAAAGAGACTTACCTTACCCGACATATA 1080

DB 1132 GTCTTTTAAATAAAGGACACAAAGAAAGAGCTTACTTACCCGACATATA 1191  
QY 1081 CCTGCGGAGATGTTCTCTCCAGTTCAATTTTACTGTGTCTTGAATCCGAGCAATTCCT 1140  
DB 1192 CCTGCGGAGATGTTCTCTCCAGTTCAATTTTACTGTGTCTTGAATCCGAGCAATTCCT 1251  
QY 1141 AAAAAGCATTTTGTGAGACCCCTTGTGACATATACAGTGAACGTGTGAAAAGCGCAG 1200  
DB 1252 AAAAAGCATTTTGTGAGACCCCTTGTGACATATACAGTGAACGTGTGAAAAGCGCAG 1311  
QY 1201 CTACAGAGCTGTGACGTAAACAGAGGAGCGAGATTAAGCGGCTTGTGACAGATGCG 1260  
DB 1312 CTACAGAGCTGTGACGTAAACAGAGGAGCGAGATTAAGCGGCTTGTGACAGATGCG 1371  
QY 1261 TGTGCTGCTTGTGATCTCTCTCGCTTCCCTTCTGCGAGCAACCACTCACTCTC 1320  
DB 1372 TGTGCTGCTTGTGATCTCTCTCGCTTCCCTTCTGCGAGCAACCACTCACTCTC 1431  
QY 1321 CTGCTGAAACATCTTCTTAACTTCAACCCAGACCATATTCGTGTGCAAGCTCAAGTTA 1380  
DB 1432 CTGCTGAAACATCTTCTTAACTTCAACCCAGACCATATTCGTGTGCAAGCTCAAGTTA 1491  
QY 1381 TTTCAACCAGAAAGGCTCAATTTGTCTTCAATGTGAAATTTCTGTACTGACACA 1440  
DB 1492 TTTCAACCAGAAAGGCTCAATTTGTCTTCAATGTGAAATTTCTGTACTGACACA 1551  
QY 1441 ACAGAGTTCTGCGAGAGGAGATATGTACAGGCTGCGCTGTGTGTGCTTCAAGTT 1500  
DB 1552 ACAGAGTTCTGCGAGAGGAGATATGTACAGGCTGCGCTGTGTGTGCTTCAAGTT 1611  
QY 1501 CTTGACCGAAACATCATGATCATCCATGAAAGACGGGGAAAGCCCTGCTCTTAAGATA 1560  
DB 1612 CTTGACCGAAACATCATGATCATCCATGAAAGACGGGGAAAGCCCTGCTCTTAAGATA 1671  
QY 1561 TCCATCTCTCTGAAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
DB 1672 TCCATCTCTCTGAAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1731  
QY 1621 ATATAGTGGGTTCAGAAACCGGACATAGCCCGTATTTTGGGTTCTTACAACTAGAGAG 1680  
DB 1732 ATATAGTGGGTTCAGAAACCGGACATAGCCCGTATTTTGGGTTCTTACAACTAGAGAG 1791  
QY 1681 AAATCTCAAGAACACACCCAGATGGAATTTTGAAGCAATGTG--GTTTTTGGCTGC 1737  
DB 1792 AAATCTCAAGAACACACCCAGATGGAATTTTGAAGCAATGTG--GTTTTTGGCTGC 1851  
QY 1738 AAGCATTAAGGATTAAGGATTAATCTATTCAGAAAAGAGCTCAGACATTTCTTAAGATGGG 1797  
DB 1852 AAGCATTAAGGATTAAGGATTAATCTATTCAGAAAAGAGCTCAGACATTTCTTAAGATGGG 1911  
QY 1798 ATCTTAATCTATTAAGGTTTCTTCTCAAGAGATGCTCTGTGGGAGAGAGAGGCC 1857  
DB 1912 ATCTTAATCTATTAAGGTTTCTTCTCAAGAGATGCTCTGTGGGAGAGAGAGGCC 1971  
QY 1858 CAGCAAGATATATACAGACACATCACTGCTTATGCGCAGAGGTGCGAGATCTCTC 1917  
DB 1972 CAGCAAGATATATACAGACACATCACTGCTTATGCGCAGAGGTGCGAGATCTCTC 2031  
QY 1918 CTTCAGAGAAAGGCGCATTTTATGTGTGAGATGCAAGAAATATGGCCAGAGATGA 1977  
DB 2032 CTTCAGAGAAAGGCGCATTTTATGTGTGAGATGCAAGAAATATGGCCAGAGATGA 2091  
QY 1978 CATGATGCCCTTGTGCAAAATATACCAAGAGGTTGAGTTGAAAATCAGAAAGCAATG 2037  
DB 2092 CATGATGCCCTTGTGCAAAATATATACCAAGAGGTTGAGTTGAAAATCAGAAAGCAATG 2151  
QY 2038 AAAACCTTGCCACTTTAAAGAGAAAAAGCTTACTTCAAGATATTTTGGTATATA 2094  
DB 2152 AAAACCTTGCCACTTTAAAGAGAAAAAGCTTACTTCAAGATATTTTGGTATATA 2208

RESULT 5

CQ726091 LOCUS 3241 bp DNA linear PAT 03-FEB-2004  
DEFINITION CQ726091 Sequence 12025 from Patent WO02068579.  
ACCESSION CQ726091  
VERSION CQ726091.1 GI:42289134  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1  
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
TITLE Kite, such as nucleic acid arrays, comprising a majority of  
humanexons or transcripts, for detecting expression and other uses  
hereof  
JOURNAL Patent: WO 02068579-A 12025 06-SEP-2002;  
PS Corporation  
FEATURES  
source 1..3241  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 83.7%; Score 1752; DB 6; Length 3241;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2092; Conservative 0; Mismatches 2; Indels 3; Gaps 1;  
QY 1 ATGAGAGGTTTCTGTTACTATATGCTACACAGACAGGACAGGCAAGCCATCGAGAA 60  
DB 80 ATGAGAGGTTTCTGTTACTATATGCTACACAGACAGGACAGGCAAGCCATCGAGAA 139  
QY 61 GAAATGTGAGAGAGCTGTGTGACATGATTTCTGCGAGATCTTCACTGTATTAAGTAA 120  
DB 140 GAAATGTGAGAGAGCTGTGTGACATGATTTCTGCGAGATCTTCACTGTATTAAGTAA 199  
QY 121 TCCGATTAAGTATGACCTTAATAAACCGAAACAGCTCCTCTGTTGTGTGTTCTACAG 180  
DB 200 TCCGATTAAGTATGACCTTAATAAACCGAAACAGCTCCTCTGTTGTGTGTTCTACAG 259  
QY 181 GGCACCGAGAGACCCAGCCAGACAGCCCGCAAGTTGTTAAGAAATACAGAACCAACA 240  
DB 260 GGCACCGAGAGACCCAGCCAGACAGCCCGCAAGTTGTTAAGAAATACAGAACCAACA 319  
QY 241 CTGCCGTTGATTTCTTTGCTCACTGCGGTATGAGGTACTGGTCTCGGTATTCAGAA 300  
DB 320 CTGCCGTTGATTTCTTTGCTCACTGCGGTATGAGGTACTGGTCTCGGTATTCAGAA 379  
QY 301 TACACCTACTTTTGAATGGGGGAGATTAATGATTAACGACTTCAAGAGCTTGAGCC 360  
DB 380 TACACCTACTTTTGAATGGGGGAGATTAATGATTAACGACTTCAAGAGCTTGAGCC 439  
QY 361 CGGCAATTTCTATGACACTGACATGACATGACTGTGTAGTTTGAACCTTGTTGAG 420  
DB 440 CGGCAATTTCTATGACACTGACATGACATGACTGTGTAGTTTGAACCTTGTTGAG 499  
QY 421 CGGTGATTTGCTGATCTGTGGCCAGCCCTTGAAAGATTTTATGTCAGACAGAGACA 480  
DB 500 CGGTGATTTGCTGATCTGTGGCCAGCCCTTGAAAGATTTTATGTCAGACAGAGACA 559  
QY 481 GAGGAGATTAAGTGGGCACTCCGGTGGGATCACTCGACCTTGTAGGACAGACTTGG 540  
DB 560 GAGGAGATTAAGTGGGCACTCCGGTGGGATCACTCGACCTTGTAGGACAGACTTGG 619  
QY 541 AAGTCAGAGCTGCTACATTTGATCTCAAGTCGACCTTCTGAGATTCGATGATTCAGAA 600  
DB 620 AAGTCAGAGCTGCTACATTTGATCTCAAGTCGACCTTCTGAGATTCGATGATTCAGAA 679  
QY 601 AGAAAGATTTCTGAGTTTGAAGCAAAATGACGTGACAGCAACCAATCCATGTTGTA 660  
DB 680 AGAAAGATTTCTGAGTTTGAAGCAAAATGACGTGACAGCAACCAATCCATGTTGTA 739  
QY 661 ATGAAGACTTGAAGTCTCACTTACCCGTTGGTACCCCACTGCAAGCCTCTG 720

DB 740 ATTGAAGACTTGAAGTCTCACTTACCCGTTGGTGAACCCCACTGCAACAGCCTCTG 799  
QY 721 AATATTCCTGTTTACCCCAAGATATTATTAAGATATCTGACAGAGCTCTGGCCAG 780  
DB 800 AATATTCCTGTTTACCCCAAGATATTATTAAGATATCTGACAGAGCTCTGGCCAG 859  
QY 781 GAGAAAGCCAGATATCTGTGACTTCAAGATTCAGATTCAGATTCAGATTCAGATTC 840  
DB 860 GAGAAAGCCAGATATCTGTGACTTCAAGATTCAGATTCAGATTCAGATTCAGATTC 919  
QY 841 GCAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 900  
DB 920 GCAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 979  
QY 901 TCAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 960  
DB 980 TCAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 1039  
QY 961 GATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 1020  
DB 1040 GATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 1099  
QY 1021 GTCTTTTGAATAAAGGACAGACAAAGAAAGAGCTACCTTACCCAGCATATA 1080  
DB 1100 GTCTTTTGAATAAAGGACAGACAAAGAAAGAGCTACCTTACCCAGCATATA 1159  
QY 1081 CCGGGGAGATTTCTCTCAAGTTCAATTTTACCTGTGCTTGAATCCGAGCAATTCCT 1140  
DB 1160 CCGGGGAGATTTCTCTCAAGTTCAATTTTACCTGTGCTTGAATCCGAGCAATTCCT 1219  
QY 1141 AAAAGGCAATTTTTCAGAGCCCTGTGAGCTATACAGTACAGTGTGAAAAGGACAG 1200  
DB 1220 AAAAGGCAATTTTTCAGAGCCCTGTGAGCTATACAGTACAGTGTGAAAAGGACAG 1279  
QY 1201 CTACAGAGCTGTGAGTAAACAAGGGGAGCCGATTAATAGCCGCTTGTGAGATGCG 1260  
DB 1280 CTACAGAGCTGTGAGTAAACAAGGGGAGCCGATTAATAGCCGCTTGTGAGATGCG 1339  
QY 1261 TGTGCTGCTTGTGATTTCTCTCTGCTTCCCTTCTGCGACAGCACTCACTGCTC 1320  
DB 1340 TGTGCTGCTTGTGATTTCTCTCTGCTTCCCTTCTGCGACAGCACTCACTGCTC 1399  
QY 1321 CTGCTGGAACATCTTCTTAACTTCAACCAACCAATATGCTGCAAGCTCAAGTTTA 1380  
DB 1400 CTGCTGGAACATCTTCTTAACTTCAACCAACCAATATGCTGCAAGCTCAAGTTTA 1459  
QY 1381 TTTCAACCAAGAAAGCTCAATTTGCTTCAACATTTGGAATTTCTGCTACTGCA 1440  
DB 1460 TTTCAACCAAGAAAGCTCAATTTGCTTCAACATTTGGAATTTCTGCTACTGCA 1519  
QY 1441 ACGAGGTTCTCGGAGAGGATATGACAGGCTGCTGCTGTTGTTGCTTCACTT 1500  
DB 1520 ACGAGGTTCTCGGAGAGGATATGACAGGCTGCTGCTGTTGTTGCTTCACTT 1579  
QY 1501 CTTCAGCCAAATATCATGCAATCCCATGAAGACAGGGGAAAGCCCTGCTCTAAGATA 1560  
DB 1580 CTTCAGCCAAATATCATGCAATCCCATGAAGACAGGGGAAAGCCCTGCTCTAAGATA 1639  
QY 1561 TCCATCTCTCTGCAACCAATTTCTTCACTTACAGATGACCCCTCAATCCCATTC 1620  
DB 1640 TCCATCTCTCTGCAACCAATTTCTTCACTTACAGATGACCCCTCAATCCCATTC 1699  
QY 1621 ATTAATGTTGTTCAAGAAACCGGATAGCCCGGTTTATGAGTTCTTACATATAGAG 1680  
DB 1700 ATTAATGTTGTTCAAGAAACCGGATAGCCCGGTTTATGAGTTCTTACATATAGAG 1759  
QY 1681 AAATCCCAAGAACCAACCCAGATGAAATTTTGAAGCAATG---GTTTGTGGCTG 1737  
DB 1760 AAATCCCAAGAACCAACCCAGATGAAATTTTGAAGCAATG---GTTTGTGGCTG 1819  
QY 1738 AGGCAATAGATAGGATTAATCTATTCAGAAAAGCTCAGACATTCCTTAAGATGAG 1797

Db 1820 AGGCAATGAGATAGGATTTATCTATTCAGAAAAAGAGCTCAGACATTTCTTAAAGCATGGG 1879  
Qy 1798 ATCTTAATCATGTAAAGGTTTCTCTCAAGAGATGCTCTGTGGGAGAGAGAGCC 1857  
Db 1880 ATCTTAATCATGTAAAGGTTTCTCTCAAGAGATGCTCTGTGGGAGAGAGAGCC 1939  
Qy 1858 CCGCAAGATATGTACAGACAACTTCAGCTTCATGCGCAGAGAGTGGGAGATCTTC 1917  
Db 1940 CCGCAAGATATGTACAGACAACTTCAGCTTCATGCGCAGAGAGTGGGAGATCTTC 1999  
Qy 1918 CTCGAGAGAGAGGCGCATATTTATGTTGTGTGAGATGAGCAAGATATGCGCAAGATGTA 1977  
Db 2000 CTCGAGAGAGAGGCGCATATTTATGTTGTGTGAGATGAGCAAGATATGCGCAAGATGTA 2059  
Qy 1978 CATGATGCGCTTGTGCAATATATAGCAAGAGTGGAGTTGAAAAACTAGAGCAATG 2037  
Db 2060 CATGATGCGCTTGTGCAATATATAGCAAGAGTGGAGTTGAAAAACTAGAGCAATG 2119  
Qy 2038 AAAACCTGGCCCTTTAAAGAAAGAAAAAGCTTACCTTCAAGATATTTGTCTATA 2094  
Db 2120 AAAACCTGGCCCTTTAAAGAAAGAAAAAGCTTACCTTCAAGATATTTGTCTATA 2176

## RESULT 6

BC054816

LOCUS

DEFINITION

Homo sapiens 5-methyltetrahydrofolate-homocysteine  
methyltransferase reductase, mRNA (CDNA clone IMAGE:5205285),  
partial cds.

BC054816.1 GI:33392775

Accession  
KEYWORDS  
SOURCE  
ORGANISM

Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 3310)

## REFERENCE

AUTHORS

Klausner, R.D., Collins, F.S., Wagner, L., Stein, C.M., Schuler, G.D.,  
Altschul, S.F., Zeeberg, B., Buettow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heib, F.,  
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
Sheets, T.E., Brownstein, M.U., Utsch, T.B., Toshiyuki, S.,  
Carninci, P., Prange, C., Raha, S., Localliano, N.A., Peters, G.J.,  
Abrams, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,  
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
Vallion, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Pahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,  
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
Butterfield, Y.S., Krzywicki, M.I., Skalske, U., Smalins, D.E.,  
Schnerch, A., Schein, J.B., Jones, S.J., and Marra, M.A.,  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

REMARK  
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgas-r@mail.nih.gov](mailto:cgas-r@mail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed By: The I.M.A.G.E. Consortium (LIML)

DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc\\_mgc@nih.gov](mailto:nisc_mgc@nih.gov)  
Akter, N., Ayala, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,  
Dietrich, N.L., Grant, S., Guan, X., Gupta, J., Haghighi, P.,  
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,  
Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,  
McDowell, J., Pearson, R., Stenrjop, S., Thomas, P.J., Touchman, J.W.,  
Tsougen, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>  
Series: IRAX Plate: 115 Row: d Column: 11  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 4505278.

## FEATURES

source

Location/Qualifiers  
1..3310  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5205285"  
/cissue\_type="lung, spleen, fetal, pooled"  
/cissue\_1b="NIH MGC\_122"  
/lab\_host="DH10B"  
/note="Vector: pCMV-Sport6"

## gene

/gene="MTRR"  
/note="synonym: MSR"  
/db\_xref="LocustID:4552"  
/db\_xref="MIM:602568"

## CDS

/codon\_start=2  
/product="MTRR protein"  
/protein\_id="AAH54816.1"  
/db\_xref="GI:33392776"  
/db\_xref="LocustID:4552"  
/db\_xref="MIM:602568"

## misc\_feature

/gene="MTRR"  
/note="flavodoxin; Region: Flavodoxin"  
/db\_xref="CD:pfam0258"

## misc\_feature

/gene="MTRR"  
/note="Cysu; Region: Sulfite reductase, alpha subunit  
(flavoprotein) (Inorganic ion transport and metabolism)"  
/db\_xref="CD:COG0369"

## ORIGIN

Query Match 81.2%; Score 1701; DB 9; Length 3310;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 2091; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

Qy 1 ATGAGAGGTTTCTGTATATATGCTACAGCAGGAGCAGGCAAGGCAATCGAGAA 60  
Db 53 ATGAGAGGTTTCTGTATATATGCTACAGCAGGAGCAGGCAAGGCAATCGAGAA 112



61 GAATGTGAGCAAGCTGTGTACATGATTTTCGAGATCTCACTGATTAAGTGA 120  
113 GAAATATGAGCAAGCTGTGTACATGATTTTCGAGATCTCACTGATTAAGTGA 172  
121 TCCGATTAAGTATGACCTTAATAAACCAGAACAGCTCTCTGTGTGTGTGTCTTACACG 180  
173 TCCGATTAAGTATGACCTTAATAAACCAGAACAGCTCTCTGTGTGTGTGTCTTACACG 232  
181 GGCACCGGAGACCCACCCGACACAGCCCGCAGAGTTGTAAAGAAATACAGAACCAACA 240  
223 GGCACCGGAGACCCACCCGACACAGCCCGCAGAGTTGTAAAGAAATACAGAACCAACA 292  
241 CCGCCGGTGAATTTCTGTGTCTCACTGCGGTATGAGTTAATGAGTTCTGAGTATTCAGA 300  
293 CCGCCGGTGAATTTCTGTGTCTCACTGCGGTATGAGTTAATGAGTTCTGAGTATTCAGA 352  
301 TACACCTACTTTTGAATGGGGGGAAGATTAATGAATTAACGACTTCAAGAGCTTGAGCC 360  
353 TACACCTACTTTTGAATGGGGGGAAGATTAATGAATTAACGACTTCAAGAGCTTGAGCC 412  
361 CCGCATTTCTATGACACTGACATGACATGACATGACATGACATGACATGACATGACATG 420  
413 CCGCATTTCTATGACACTGACATGACATGACATGACATGACATGACATGACATGACATG 472  
421 CCGTGAATTTCTGACACTGCGCCAGCCCTCAGAAAACATTTTATGATGACAGAGACA 480  
473 CCGTGAATTTCTGACACTGCGCCAGCCCTCAGAAAACATTTTATGATGACAGAGACA 532  
481 GAGAGATTAAGTGGGCACTCCCGGTGATGACCTGATCTTGAAGACAGACCTTGTG 540  
533 GAGAGATTAAGTGGGCACTCCCGGTGATGACCTGATCTTGAAGACAGACCTTGTG 592  
541 AAGTCAGAGCTCTACACATTTGATCTCAAGTCAGACCTTCTGATTCAGATTCAGAGA 600  
593 AAGTCAGAGCTCTACACATTTGATCTCAAGTCAGACCTTCTGATTCAGATTCAGAGA 652  
601 AGAAAGATTTCTGAGGTTTGAAGCAAAATGACATGACAGCAACCAATCCATGTTGTA 660  
653 AGAAAGATTTCTGAGGTTTGAAGCAAAATGACATGACAGCAACCAATCCATGTTGTA 712  
661 ATTGAAGATTTGAGGCTCTCACTTACCCGTTGGGTACCCCACTCTCAAGACCTTCTG 720  
713 ATTGAAGATTTGAGGCTCTCACTTACCCGTTGGGTACCCCACTCTCAAGACCTTCTG 772  
721 AATATTTCTGTGTTTACCCCAAGATTTTACAGATCATCTGACAGAGTCTCTGCGCAG 780  
773 AATATTTCTGTGTTTACCCCAAGATTTTACAGATCATCTGACAGAGTCTCTGCGCAG 832  
781 GAGGAAAGCCAGATCTGTGACTTCAAGATCCAGTTTTCAGTGCCTTAATTCAGAG 840  
833 GAGGAAAGCCAGATCTGTGACTTCAAGATCCAGTTTTCAGTGCCTTAATTCAGAG 892  
841 GAGGATCACTTATCTAGATGATGCTCATTAACCACTCTGCTGATTAATTCAGAG 900  
893 GAGGATCACTTATCTAGATGATGCTCATTAACCACTCTGCTGATTAATTCAGAG 952  
901 TCAATATCAAGACTTTTCTATCAGCTTGAAGATGCTTCAAGGCTGATCTGCTTACAGT 960  
953 TCAATATCAAGACTTTTCTATCAGCTTGAAGATGCTTCAAGGCTGATCTGCTTACAGT 1012  
961 GATTTGAGGTACAAAGCTTATCTCAAGATCTGAGCTTGAAGATTAAGAGAGCCTGC 1020  
1013 GATTTGAGGTACAAAGCTTATCTCAAGATCTGAGCTTGAAGATTAAGAGAGCCTGC 1072  
1021 GTCTTTTGAATAAAGGACACAAAGAAAGAGCTTACTTACCCAGCATATA 1080  
1073 GTCTTTTGAATAAAGGACACAAAGAAAGAGCTTACTTACCCAGCATATA 1132  
1081 CCGGCGGAGATGTTCTCTCAGTTCAATTTTCTGAGTCTTGAATAATCCAGAGCAATTCCT 1140  
1133 CCGGCGGAGATGTTCTCTCAGTTCAATTTTCTGAGTCTTGAATAATCCAGAGCAATTCCT 1192

1141 AAAAGCATTTTTCGAGCCCTTGTGACATTAACAGTGAACAGTCTGAAGAGCCAGG 1200  
1193 AAAAGCATTTTTCGAGCCCTTGTGACATTAACAGTGAACAGTCTGAAGAGCCAGG 1252  
1201 CTACAGAGCTGTGACATTAACAGAGGAGCCGATTAATAGCCGTTTGTACAGATGCC 1260  
1253 CTACAGAGCTGTGACATTAACAGAGGAGCCGATTAATAGCCGTTTGTACAGATGCC 1312  
1261 TGTGCTGCTTGTGATCTCTCTGCTGCTTCCCTTCTGACAGCCACACTGATCTC 1320  
1313 TGTGCTGCTTGTGATCTCTCTGCTGCTTCCCTTCTGACAGCCACACTGATCTC 1372  
1321 CTGCTGCAACATCTCTTAACCTTCAACAGACCATTAATGCTGTGACAGTCAAGTTTA 1380  
1373 CTGCTGCAACATCTCTTAACCTTCAACAGACCATTAATGCTGTGACAGTCAAGTTTA 1432  
1381 TTTCAACCAAGAAAGCTCATTTTGTCTTCAACATTTGAAATTTTCTGTCTACCTGCCA 1440  
1433 TTTCAACCAAGAAAGCTCATTTTGTCTTCAACATTTGAAATTTTCTGTCTACCTGCCA 1492  
1441 ACAGAGTTTCTGCGGAAAGGATATGACAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
1493 ACAGAGTTTCTGCGGAAAGGATATGACAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTG 1552  
1501 CTTGACGCAACATCATGATCCCATGAAGACAGCGGAAAGCCCTGCTCTTAAGATA 1560  
1553 CTTGACGCAACATCATGATCCCATGAAGACAGCGGAAAGCCCTGCTCTTAAGATA 1612  
1561 TCCATCTCTCTGCAACAAATTTCTTCACTTACAGATGACCCCTGAATCCCATC 1620  
1613 TCCATCTCTCTGCAACAAATTTCTTCACTTACAGATGACCCCTGAATCCCATC 1672  
1621 ATTAATGTTGGTTCAGAAACCGGATAGCCCGTTTATGGTTCCTTCAACATGAGAG 1680  
1673 ATTAATGTTGGTTCAGAAACCGGATAGCCCGTTTATGGTTCCTTCAACATGAGAG 1732  
1681 AAATCCAGAAACAACACCCAGATGAATTTTGAAGCAATGTG--GTTTGGCTGC 1737  
1733 AAATCCAGAAACAACACCCAGATGAATTTTGAAGCAATGTG--GTTTGGCTGC 1792  
1738 AGGATTAAGATGAGATTAATCTATTCAGAAAAGACTCAGACATTTCTTAAAGATGG 1797  
1793 AGGATTAAGATGAGATTAATCTATTCAGAAAAGACTCAGACATTTCTTAAAGATGG 1852  
1798 ATCTTAATCTATTAAGGTTTCTTCTCAGAGATGCTCTCTGTTGGGAGAGAGAGCC 1857  
1853 ATCTTAATCTATTAAGGTTTCTTCTCAGAGATGCTCTCTGTTGGGAGAGAGAGCC 1912  
1858 CCAGCAAGATATGTAACAAGAACATCCAGCTTCAATGCGCAGAGTGGCAGAAATCTCTC 1917  
1913 CCAGCAAGATATGTAACAAGAACATCCAGCTTCAATGCGCAGAGTGGCAGAAATCTCTC 1972  
1918 CTCCAGAGAAAGGCAATTAATTAATGTTGTGTGAGATGCAAGAAATATGCAAGATGTA 1977  
1973 CTCCAGAGAAAGGCAATTAATTAATGTTGTGTGAGATGCAAGAAATATGCAAGATGTA 2032  
1978 CATGATGCTTGTGCAATTAATTAAGCAAGAGTTTGAAGTTGAATAACTAGAAGCAATG 2037  
2033 CATGATGCTTGTGCAATTAATTAAGCAAGAGTTTGAAGTTGAATAACTAGAAGCAATG 2092  
2038 AAAACCTGGGCACTTTAAAGAAAGAAAGCTTACAGATATTTGATCATTA 2094  
2093 AAAACCTGGGCACTTTAAAGAAAGAAAGCTTACAGATATTTGATCATTA 2149

RESULT 7  
BVL77620/c 2933 bp DNA linear STS 10-JUN-2004  
LOCUS BVL77620 Human DNA (Sequenc) Homo sapiens STS genomic, Sequence  
DEFINITION tagged site.  
ACCESSION BVL77620  
VERSION BVL77620.1 GI:48013757  
KEYWORDS STS.



SOURCE	Homo sapiens (human)
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE	1 (bases 1 to 2933)
AUTHORS	Nelson,R.M., Marnellio,G., Kammerer,S., Hoyal,C.R., Shi,M.M., Cantor,C.R. and Braun,A.
TITLE	Large-Scale Validation of Single Nucleotide Polymorphisms in Gene Regions
JOURNAL	Genome Res. (2004) In press
COMMENT	Contact: Andreas Braun Pharmaceuticals division Sequenom, Inc. 3595 John Hopkins Court, San Diego, CA 92121, USA Tel: 18582028018 Fax: 18582023020 Email: abraun@sequenom.com Primer A: No primer sequence submitted Primer B: No primer sequence submitted STS size: 2933.
FEATURES	Location/Qualifiers
source	1..2933 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /clone_1fb="Human DNA (Sequenom)" <1..>2933
ORIGIN	STS
Query Match	58.3%; Score 1220; DB 11; Length 2933;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 1220; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Oy	315 CAATGGGGGGAAGAATTGTAATAACGACTTCAAGAGCTGGAGCCGGGCAATTTCTATGA 374
Db	2846 CAAATGGGGGGAAGATTAATTGATTAACGACTTCACAAGCCTTGAGGCCGGCATTTTCTATGA 2788
Oy	375 CACTGGACATGCAGATGA CTGTGTAAGTTTAGAACCTTGTTGAGCCGTGGATTCTGG 434
Db	2786 CACTGGA CATGCAGATGACTGTGTAGTTTAGAACCTTGTTGAGCCGTGGATTCTGG 2722
Oy	435 ACTGTGGCACCGCTCABAAGCAATTTNAGTCAGACGAGACAGACAAAGACGATTAATGG 494
Db	2726 ACTGTGGCAGCGCTCABAAGCAATTTAAGTCAGACGAGACAGACAAAGAGATTAATGG 2667
Oy	495 CGCACTCCCGGTGGCATCACTGCATCTCTTGAGAGACAGACCTTGTGAAGTCAGAGCTCT 554
Db	2666 CGCACTCCCGGTGGCATCACTGCATCTCTTGAGAGACAGACCTTGTGAAGTCAGAGCTCT 2607
Oy	555 ACAATTGAATCTCAAGTCGAGCTTCTGAGATTGATGATTCAGAGAAAGAAATTTCTGA 614
Db	2606 ACACATTAATCTCAAGTCGAGCTTCTGAGATTGATGATTCAGAGAAAGAAATTTCTGA 2547
Oy	615 GGTTTTGAAGCAAATGACAGTGAAGAAACCAATCCAAATGNTTAATTGAAGACTTTGA 674
Db	2546 GGTTTTGAAGCAAATGACAGTGAAGAAACCAATCCAAATGNTTAATTGAAGACTTTGA 2487
Oy	675 GTCTCACTTACCCTGTCGGTACCCCACTCTCACAGCTCTCTGAATATTCTCTGGTT 734
Db	2486 GTCTCACTTACCCTGTCGGTACCCCACTCTCACAGCTCTCTGAATATTCTCTGGTT 722
Oy	735 ACCCCCAAAATTTTACAGGTCATCTGCAAGAGTCTCTTGGCCAGAGAAAGCAAGT 794
Db	2426 ACCCCCAAAATTTTACAGGTCATCTGCAAGAGTCTCTTGGCCAGAGAAAGCAAGT 2367
Oy	795 ATCTGTGACTTCAGAGATCCAGTTTTTCAATGTCAAATTTCAAGAGAGTTCAACTTAC 854
Db	2366 ATCTGTGACTTCAGAGATCCAGTTTTTCAATGTCAAATTTCAAGAGAGTTCAACTTAC 2307
Oy	855 TACGATGATGCCAATAAAAACCACTCTGCTGTGTAGAAATTTGACATTTCAATACGACTT 914
Db	2306 TACGATGATGCCAATAAAAACCACTCTGCTGTGTAGAAATTTGACATTTCAATACGACTT 2247

Oy	915	TTCCCTATCAGCCCTGAGAAATGCGCTTACAGCGATCTGTGCCCTTAACAGTAACTCTGAGGTACA	974
Db	2246	TTCCCTATCAGCCCTGAGAAATGCGCTTACAGCGATCTGTGCCCTTAACAGTAACTCTGAGGTACA	2187
Oy	975	AAGCTACTCCAAAGACTGAGCTTGGAAGATTAAGAAGACACTGCTCTTTTGAAAT	1034
Db	2186	AAGCTACTCCAAAGACTGAGCTTGGAAGATTAAGAAGACACTGCTCTTTTGAAAT	2127
Oy	1035	AAAGGCGACACAAAGAAAGAAAGAGCTACTTAACCCAGCATTAACCTGGGAGTGTTC	1094
Db	2126	AAAGGCGACACAAAGAAAGAAAGAGCTACTTAACCCAGCATTAACCTGGGAGTGTTC	2067
Oy	1095	TCTCAGCTTCAATTTTAACTGCTGTCTTGAATCCAGACATTCCTAAAGGCAATTTT	1154
Db	2066	TCTCAGCTTCAATTTTAACTGCTGTCTTGAATCCAGACATTCCTAAAGGCAATTTT	2007
Oy	1155	GCGAGCCCTTGAGACTATACCACTGACAGTGTGAAGAGCGACGCTACAGAGCTGTG	1214
Db	2006	GCGAGCCCTTGAGACTATACCACTGACAGTGTGAAGAGCGACGCTACAGAGCTGTG	1947
Oy	1215	CAGTAAACAAAGGCGACGCCGATTAATGACCGCTTTGTATCGAGATGCTGTGCTTGT	1274
Db	1946	CAGTAAACAAAGGCGACGCCGATTAATGACCGCTTTGTATCGAGATGCTGTGCTTGT	1887
Oy	1275	GGATCTCTCCTCGCTTCCCTTTGTCGACGACCACTAGTCTCCTGCTCGAACATCT	1334
Db	1886	GGATCTCTCCTCGCTTCCCTTTGTCGACGACCACTAGTCTCCTGCTCGAACATCT	1827
Oy	1335	TCCCTAACTTCAACCCAGACCATATTCGTGTGCAAGCTCAAGTTATTTACCCAGGAAA	1394
Db	1826	TCCCTAACTTCAACCCAGACCATATTCGTGTGCAAGCTCAAGTTATTTACCCAGGAAA	1767
Oy	1395	GCTTCATTTTGTCTTCAACATTTGTGAAATTTCTGTCTACCTGCCACACAGAGTTCTGCG	1454
Db	1766	GCTTCATTTTGTCTTCAACATTTGTGAAATTTCTGTCTACCTGCCACACAGAGTTCTGCG	1707
Oy	1455	GAAGGAGATATGTACAGAGCTGCGCTGGCTTTGGTTGCTTCACTTCAGTCTTCGACCAACAT	1514
Db	1706	GAAGGAGATATGTACAGAGCTGCGCTGGCTTTGGTTGCTTCACTTCAGTCTTCGACCAACAT	1647
Oy	1515	ACATGCATCCCATGAGACA	1534
Db	1646	ACATGCATCCCATGAGACA	1627
RESULT 8			
BV178010/c			
LOCUS	sqm97986 Human DNA (Sequenom) Homo sapiens STS genomic, sequence	2933 bp	linear STS 10-JUN-2004
DEFINITION	tagged site.		
ACCESSION	BV178010		
VERSION	BV178010.1	GI:48014252	
KEYWORDS	STS.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 2933)		
JOURNAL	NeJson,R.M., Mannellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,		
COMMENT	Canter,C.R. and Braun,A.		
	Large-Scale Validation of Single Nucleotide Polymorphisms in Gene		
	Regions		
	Genome Res. (2004) In press		
	Contact: Andreas Braun		
	Pharmaceuticals division		
	Sequenom, Inc.		
	3595 John Hopkins Court, San Diego, CA 92121, USA		
	Tel.: 18582028018		
	Fax: 18582028020		
	Email: abraun@sequenom.com		
	Primer A: No primer sequence submitted		



Db 63 AAAATAAGGACAGACAAAGAGAGAGCTACCTTACCACGATATACCTGGGGA 122  
QY 1090 TGTCTCTCCAGTTCATTTTCTGCTGTCTTGAATCCGAGCAATTCCTAAAAAGCA 1149  
Db 123 TGTCTCTCCAGTTCATTTTCTGCTGTCTTGAATCCGAGCAATTCCTAAAAAGCA 182  
QY 1150 TTTTGGAGAGCCCTGTGGAGTATACAGAGAGAGCTGTAAGAAAGCCAGGCTACAGAG 1209  
Db 183 TTTTGGAGAGCCCTGTGGAGTATACAGAGAGAGCTGTAAGAAAGCCAGGCTACAGAG 242  
QY 1210 CTGTGAGTAAAGAGAGAGAGAGAGAGTATAGCCGCTTTGAGAGATGCTGTGCTGC 1269  
Db 243 CTGTGAGTAAAGAGAGAGAGAGAGAGTATAGCCGCTTTGAGAGATGCTGTGCTGC 302  
QY 1270 TTGTGGATCT 1329  
Db 303 TTGTGGATCT 362  
QY 1330 CATCTTCTTAACTTCAACCCAGAC 1355  
Db 363 CATCTTCTTAACTTCAACCCAGAC 388

RESULT 10  
LOCUS F121202S04 1353 bp DNA linear PRI 14-DEC-1999  
DEFINITION Homo sapiens methionine synthase reductase (MTRR) gene, exon 5.  
ACCESSION AF121205 GI:6572530  
VERSION AF121205.1 GI:6572530  
KEYWORDS 4 of 12  
SEGMENT Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1353)  
AUTHORS Leclerc, D., Odievre, M., Wu, Q., Wilson, A., Hutzenga, J.J., Rozen, R.,  
Scheer, S.W. and Gravel, R.A.  
TITLE Molecular cloning, expression and physical mapping of the human  
methionine synthase reductase gene  
JOURNAL Gene 240 (1), 75-88 (1999)  
MEDLINE 20033550  
PUBMED 10564814  
REFERENCE 2 (bases 1 to 1353)  
AUTHORS Leclerc, D.  
TITLE Direct Submision  
COMMENT Submitted (20-JAN-1999) Human Genetics, Montreal Children's  
Hospital, 4060 Ste-Catherine West, Montreal, Quebec H3Z 2Z3, Canada  
LOCATION/Qualifiers  
FEATURES  
source 1..1353  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/map="5p15.2-p15.3; between W11755 and D5S1957"  
exon 358..736  
/gene="MTRR"  
/number=5

## ORIGIN

Query Match 18.2%; Score 381; DB 9; Length 1353;  
Best Local Similarity 100.0%; Pred. No. 1.8e-196;  
Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 401 GTTTAGAACTTGTGTGAGCCGTGATGCTGAGACTGTGGCAGCCCTCAGAAACATT 460  
Db 357 GTTTAGAACTTGTGTGAGCCGTGATGCTGAGACTGTGGCAGCCCTCAGAAACATT 416  
QY 461 TTAGGTCAAGCAGAGACAGAGAGATTAAGTGGCGCACTCCGGTGGCATCACTGCAT 520  
Db 417 TTAGGTCAAGCAGAGACAGAGAGATTAAGTGGCGCACTCCGGTGGCATCACTGCAT 476  
QY 521 CCTTGAAGAGACCTTGTGAAGTCAAGCTGCTACACATTGAATCTCAAGTGGAGCTTC 580

Db 477 CTTGAGACAGACCTTGTGAAGTCAAGCTGTACACATTGAATCTCAAGTGGAGCTTC 536  
QY 581 TGAATTCGATGATTCAGAGAGAGAGATTCAGAGGTTTGAAGCAAAATGCAGTACAG 640  
Db 537 TGAATTCGATGATTCAGAGAGAGAGATTCAGAGGTTTGAAGCAAAATGCAGTACAG 596  
QY 641 GCAACCAATCCAAATGTTGTAATTGAAGACTTTGAGTCTCACTTACCCTGGTACCCC 700  
Db 597 GCAACCAATCCAAATGTTGTAATTGAAGACTTTGAGTCTCACTTACCCTGGTACCCC 656  
QY 701 CACTTCACAAGCCTCTGTAATATTCCTGTTTACCCCAAGATATTTACAGAGTATC 760  
Db 657 CACTTCACAAGCCTCTCTGAATATTCCTGTTTACCCCAAGATATTTACAGAGTATC 716  
QY 761 TGCAGAGTCTCTTGGCCAGG 781  
Db 717 TGCAGAGTCTCTTGGCCAGG 737

RESULT 11  
LOCUS AC010346 109626 bp DNA linear PRI 10-NOV-2000  
DEFINITION Homo sapiens chromosome 5 clone C17B-H1\_201B2, complete sequence.  
ACCESSION AC010346  
VERSION AC010346.6 GI:11136705  
KEYWORDS HTG.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 109626)  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submision  
COMMENT Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
JOURNAL 3 (bases 1 to 109626)  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submision  
COMMENT Submitted (10-NOV-2000) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
On Nov 10, 2000 this sequence version replaced gi:9256196.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www-shgc.stanford.edu  
Quality: Phrap Quality >=40 99.9% of Sequence;  
STS Content:  
WI-9255 G05749.

FEATURES  
source 1..109626  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/clone="C17B-H1\_201B2"

## ORIGIN

Query Match 15.8%; Score 330; DB 9; Length 109626;  
Best Local Similarity 99.7%; Pred. No. 1.1e-168;  
Matches 380; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 401 GTTTAGAACTTGTGTGAGCCGTGATGCTGAGACTGTGGCAGCCCTCAGAAACATT 460  
Db 88571 GTTTAGAACTTGTGTGAGCCGTGATGCTGAGACTGTGGCAGCCCTCAGAAACATT 88630  
QY 461 TTAGGTCAAGCAGAGACAGAGAGATTAAGTGGCGCACTCCGGTGGCATCACTGCAT 520  
Db 88631 TTAGGTCAAGCAGAGACAGAGAGATTAAGTGGCGCACTCCGGTGGCATCACTGCAT 88690

QY 521 CCTTGAGACAGACCTTGTGAAGTCAGAGCTGCTACATTTGAATCTCAAGTCGACCTTC 580  
 DB 88691 CCTGAGAGACAGACCTTGTGAAGTCAGAGCTGCTACATTTGAATCTCAAGTCGACCTTC 88750  
 QY 581 TGAATTCGATGATTCAGAGAGAGAGATTTGAGAGCTTTTGAAGCAAAATGCAAGTGAACA 640  
 DB 88751 TGAATTCGATGATTCAGAGAGAGAGATTTGAGAGCTTTTGAAGCAAAATGCAAGTGAACA 88810  
 QY 641 GCAACCAATCCATGTTGTAATTTGAAGACTTTGAGTCTCCTACTTACCCTGCGGTATCCCC 700  
 DB 88811 GCAACCAATCCATGTTGTAATTTGAAGACTTTGAGTCTCCTACTTACCCTGCGGTATCCCC 88870  
 QY 701 CACTCTCACAAGCTCTCTGTAATATTCCTGTTTACCCCAAGAAATTTTACAGGTACATC 760  
 DB 88871 CACTCTCACAAGCTCTCTGTAATATTCCTGTTTACCCCAAGAAATTTTACAGGTACATC 88930  
 QY 761 TGCAGAGATCTCTTGGCCAGG 781  
 DB 88931 TGCAGAGATCTCTTGGCCAGG 88951  
 RESULT 12  
 AC025174 110756 bp DNA linear PRI 28-MAR-2002  
 LOCUS Homo sapiens chromosome 5 clone CTD-2072124, complete sequence.  
 DEFINITION AC025174  
 AC025174  
 VERSION AC025174.5 GI:19774456  
 KEYWORDS HTG.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 110756)  
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 110756)  
 AUTHORS DOE Joint Genome Institute.  
 JOURNAL Direct Submission  
 REFERENCE 3 (bases 1 to 110756)  
 AUTHORS DOE Joint Genome Institute.  
 JOURNAL Direct Submission  
 REFERENCE 4 (bases 1 to 110756)  
 AUTHORS DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
 JOURNAL Submitted (07-MAR-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
 COMMENT On Mar 28, 2002 this sequence version replaced gi:19224767.  
 Draft Sequence Produced by DOE Joint Genome Institute  
 www.fgi.doe.gov  
 Finishing Completed at Stanford Human Genome Center  
 www.hgsc.stanford.edu  
 Quality: Phrap Quality >=40 100% of Sequence;  
 Estimated Total Number of Errors is 0.  
 Location/Qualifiers  
 1..110756  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="5"  
 /clone="CTD-2072124"  
 ORIGIN  
 Query Match 15.8%; Score 330; DB 9; Length 110756;  
 Best Local Similarity 99.7%; Pred. No. 1,1e-168;  
 Matches 380; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 401 GTTTAGAACTTGTGTTGAGCCGTGATGCTGCACTCTGCGCAGCCCTCAGAAAGCAATT 460

DB 20040 GTTTAGAACTTGTGTTGAGCCGTGATTTGCTGCACTCTGCGCAGCCCTCAGAAAGCAATT 20099  
 QY 461 TTAGGTCAAGCAGAGAGACAGAGAGATGAGTGGGCACTCCCGGGGATCACCCTGCAT 520  
 DB 20100 TTAGGTCAAGCAGAGAGACAGAGAGATGAGTGGGCACTCCCGGGGATCACCCTGCAT 20159  
 QY 521 CCTTGAGACAGACCTTGTGAAGTCAGAGCTGCTACATTTGAATCTCAAGTCGACCTTC 580  
 DB 20160 CCTGAGAGACAGACCTTGTGAAGTCAGAGCTGCTACATTTGAATCTCAAGTCGACCTTC 20219  
 QY 581 TGAATTCGATGATTCAGAGAGAGAGATTTGAGAGCTTTTGAAGCAAAATGCAAGTGAACA 640  
 DB 20220 TGAATTCGATGATTCAGAGAGAGAGATTTGAGAGCTTTTGAAGCAAAATGCAAGTGAACA 20279  
 QY 641 GCAACCAATCCATGTTGTAATTTGAAGACTTTGAGTCTCCTACTTACCCTGCGGTATCCCC 700  
 DB 20280 GCAACCAATCCATGTTGTAATTTGAAGACTTTGAGTCTCCTACTTACCCTGCGGTATCCCC 20339  
 QY 701 CACTCTCACAAGCTCTCTGTAATATTCCTGTTTACCCCAAGAAATTTTACAGGTACATC 760  
 DB 20340 CACTCTCACAAGCTCTCTGTAATATTCCTGTTTACCCCAAGAAATTTTACAGGTACATC 20399  
 QY 761 TGCAGAGATCTCTTGGCCAGG 781  
 DB 20400 TGCAGAGATCTCTTGGCCAGG 20420  
 RESULT 13  
 AC022921 158199 bp DNA linear HTG 12-MAR-2000  
 LOCUS Homo sapiens clone RP11-138P20, WORKING DRAFT SEQUENCE, 12  
 DEFINITION AC022921  
 AC022921  
 VERSION AC022921.2 GI:7229868  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 158199)  
 AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 158199)  
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Becker, R., Beda, F., Boguslavsky, L., Boukhalter, B., Brown, A., Burkett, G., Castle, A., Choquel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearlano, K., Dewar, K., Domino, M., Doyle, M., Penestor, J., Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Landers, T., Lehoczek, J., Levine, R., Liu, C., Liu, G., Locke, K., MacDonald, P., Margulis, N., McEwan, P., McGurk, A., McKernan, K., McPherson, R., Meldrum, J., Menus, L., Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., Olivari, T.M., Peterson, K., Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rothman, D., Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testfaye, S., Theodore, J., Tirelli, A., Vassiliou, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W., J., Zimmer, A. and Zody, M.  
 JOURNAL Direct Submission  
 Submitted (07-FEB-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Mar 12, 2000 this sequence version replaced gi:6921909.  
 All repeats were identified using RepeatMasker:  
 Smit, A.P.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/XM/RepeatMasker.html  
 ----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seq.wi.mit.edu

1	1284	1283: contig of 1283 bp in length
*	1284	1283: gap of 100 bp
*	1384	4203: contig of 2820 bp in length
*	4204	4303: gap of 100 bp
*	4304	6786: contig of 2483 bp in length
*	6787	6886: gap of 100 bp
*	6887	9683: contig of 2797 bp in length
*	9684	9783: gap of 100 bp
*	9784	12902: contig of 3119 bp in length
*	12903	13002: gap of 100 bp
*	13003	16429: contig of 3427 bp in length
*	16430	16529: gap of 100 bp
*	16530	25201: contig of 8672 bp in length
*	25202	25301: gap of 100 bp
*	25302	36759: contig of 11458 bp in length
*	36760	36859: gap of 100 bp
*	36860	53921: contig of 17062 bp in length
*	53922	54021: gap of 100 bp
*	54022	72054: contig of 18033 bp in length
*	72055	72154: gap of 100 bp
*	72155	102527: contig of 30373 bp in length
*	102528	102627: gap of 100 bp
*	102628	158199: contig of 55572 bp in length.
FEATURES		
source		
1.158199		
/organism="Homo sapiens"		
/mol_type="genomic DNA"		
/db_xref="taxon:9606"		
/clone="RP11-138P20"		
/clone_lib="RP11-11 Human Male BAC"		
1.1283		
/note="assembly_fragment"		
1384. .4203		
/note="assembly_fragment"		
4304. .6786		
/note="assembly_fragment"		
6887. .9683		
/note="assembly_fragment"		
clone end:SP6		
vector side:left"		
9784. .12902		
/note="assembly_fragment"		
13003. .16429		
/note="assembly_fragment"		
16530. .25201		
/note="assembly_fragment"		
clone end:T7		
vector_side:right"		
25302. .36759		

misc_feature	/note="assembly_fragment"	36860..53921	
misc_feature	/note="assembly_fragment"	54022..72054	
misc_feature	/note="assembly_fragment"	72155..102527	
misc_feature	/note="assembly_fragment"	102628..158199	
misc_feature	/note="assembly_fragment"		
Query Match	13.3%; Score 279; DB 2; Length 158199;		
Best Local Similarity	99.5%; Pred. No. 9.6e-141;		
Matches 379; Conservative	0; Mismatches 2; Indels 0; Gaps 0;		
Db	401	GTTTAGAACTTGTGGTTGAGCCGCTGAGATTGCTGAGACTCTTGSCAGACCCCTCAGAAAGCATTT	460
Db	3446	GTTTAGAACTTGTGGTTGAGCCGCTGAGATTGCTGAGACTCTTGSCAGACCCCTCAGAAAGCATTT	3505
Qy	461	TTAGGTCAAGCAGAGGACAAAGAGAGATTAAGTGCGGCACTCCCGGTGGCATTCACCTTGAT	520
Db	3506	TTAGGTCAAGCAGAGGACAAAGAGAGATTAAGTGCGGCACTCCCGGTGGCATTCACCTTGAT	3565
Qy	521	CCTTGAGAGACAGACCTTGTGTAAGTCAGAGCGCTGACATTCATGTAATCTCAAGTGAGCTTC	580
Db	3566	CCTTGAGAGACAGACCTTGTGTAAGTCAGAGCGCTGACATTCATGTAATCTCAAGTGAGCTTC	3625
Qy	581	TGAGATTGATGATTTGAGAAAGAAAGATTTCTGAGTTTGAAGCAAAATGACAGTGACA	640
Db	3626	TGAGATTGATGATTTGAGAAAGAAAGATTTCTGAGTTTGAAGCAAAATGACAGTGACA	3685
Qy	641	GCAACCAATCCATATGTTGTATTTGAAGAATTTGAGTCTTCACTTACCCGTTGGTACCC	700
Db	3686	GCAACCAATCCATATGTTGTATTTGAAGAATTTGAGTCTTCACTTACCCGTTGGTACCC	3745
Qy	701	CACCTCAACAAGCTCTGTAATTTCTGCTGTTTACCCGAGATATTTACAGGTACATC	760
Db	3746	CACCTCAACAAGCTCTCTGTGATATTTCTGCTGTTTACCCGAGATATTTACAGGTACATC	3805
Qy	761	TGCAGAGATCTCTTGCCAGG 781	
Db	3806	TGCAGAGATCTCTTGCCAGG 3826	
RESULT 14			
AC021609	167237 bp	DNA	linear HTG 12-MAR-2000
LOCUS			
DEFINITION	Homo sapiens clone RP11-259D10, WORKING DRAFT SEQUENCE, 6 unordered		
ACCESSION	AC021609		
VERSION	AC021609.3		
KEYWORDS	HTG; HTGS PHASE1; HTGS_DRAFT.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Bukharin, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 167237)		
JOURNAL	Bitren, B., Linton, L., Nusbaum, C. and Lander, E.		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 167237)		
	Bitren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,		
	Anderson, S., Baldwin, J., Barina, N., Beckert, R., Bede, F.,		
	Boguslavsky, I., Boukhalter, B., Brown, A., Burkett, G., Castle, A.,		
	Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,		
	DeArrelano, K., Dewar, K., Domino, M., Doyle, M., Fenebor, J.,		
	Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J.,		
	Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,		
	Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatae, A., Klein, J.,		
	Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G., Locke, K.,		
	Macdonald, P., Margulis, N., McEwan, P., McGuck, A., McKernan, K.,		
	McHeaters, R., Meldrim, J., Meeneus, L., Morrow, J., Naylor, J.,		
	Norman, C. H., O'Connor, T., O'Donnell, P., Oliver, T. M., Peterson, K.,		



Quality coverage: 10.01 in Q20 bases; sum-of-contigs estimation.

\* NOTE: This is a 'working draft' sequence. It currently consists of 27 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```
1 1207: contig of 1207 bp in length
* 1208 1307: gap of unknown length
* 1308 3385: contig of 2078 bp in length
* 3386 3485: gap of unknown length
* 3486 5553: contig of 2068 bp in length
* 5554 5653: gap of unknown length
* 5654 7030: contig of 1377 bp in length
* 7031 7131: gap of unknown length
* 7131 9069: contig of 1939 bp in length
* 9070 9169: gap of unknown length
* 9170 11452: contig of 2283 bp in length
* 11453 11553: gap of unknown length
* 11553 16002: contig of 4450 bp in length
* 16003 16102: gap of unknown length
* 16103 21083: contig of 4981 bp in length
* 21084 2183: gap of unknown length
* 2184 26893: contig of 5710 bp in length
* 26894 26993: gap of unknown length
* 26994 29987: contig of 2994 bp in length
* 29988 30087: gap of unknown length
* 30088 32949: contig of 2862 bp in length
* 32950 33049: gap of unknown length
* 33050 38757: contig of 5708 bp in length
* 38758 38857: gap of unknown length
* 38858 45202: contig of 6345 bp in length
* 45203 45302: gap of unknown length
* 45303 50911: contig of 5609 bp in length
* 50912 51011: gap of unknown length
* 51012 55880: contig of 4869 bp in length
* 55881 55980: gap of unknown length
* 55981 63189: contig of 7209 bp in length
* 63190 63289: gap of unknown length
* 63290 69632: contig of 6343 bp in length
* 69633 69732: gap of unknown length
* 69733 76378: contig of 6645 bp in length
* 76379 76477: gap of unknown length
* 76478 83051: contig of 6574 bp in length
* 83052 83151: gap of unknown length
* 83152 94929: contig of 11778 bp in length
* 94930 95029: gap of unknown length
* 95030 104346: contig of 9317 bp in length
* 104347 104446: gap of unknown length
* 104447 115275: contig of 10829 bp in length
* 115276 115375: gap of unknown length
* 115376 123488: contig of 8113 bp in length
* 123489 123588: gap of unknown length
* 123589 133172: contig of 9584 bp in length
* 133173 133272: gap of unknown length
* 133273 145495: contig of 12223 bp in length
* 145496 145595: gap of unknown length
* 145596 160270: contig of 14675 bp in length
* 160271 160370: gap of unknown length
* 160371 177596: contig of 17226 bp in length.
```

## FEATURES

source

```
1. 177596
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="RP11-35616"
/location="RP11 human BAC library 11"
```

## ORIGIN

Query Match

13.3%; Score 279; DB 2; Length 177596;

```
Best Local Similarity 99.5%; Pred. No. 9.5e-141;
Matches 379; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 401 GTTTAGAACTTGTGGTTGAGCCGTGATTTGCGACTGCGCAGCCCTCAGAAAGCATT 460
Db 83524 GTTTAGAACTTGTGGTTGAGCCGTGATTTGCGACTGCGCAGCCCTCAGAAAGCATT 83583
QY 461 TTAGGTCAAGCAGAGACAAAGAGATTAAGTGGCGCACTCCCGGTGGATCACTGCAAT 520
Db 83584 TTAGGTCAAGCAGAGACAAAGAGATTAAGTGGCGCACTCCCGGTGGATCACTGCAAT 83643
QY 521 CCTTAGAGCAGACCTTGTGAAGTCAAGCTGCTACATTAATCAAGTCAGACTTC 580
Db 83644 CCTTAGAGCAGACCTTGTGAAGTCAAGCTGCTACATTAATCAAGTCAGACTTC 83703
QY 581 TGAGATTCAGATTCAGAGAAAGAAAGATTTGAGTTTGAAGCAAAATGAGTGAACA 640
Db 83704 TGAGATTCAGATTCAGAGAAAGAAAGATTTGAGTTTGAAGCAAAATGAGTGAACA 83763
QY 641 GCAACCAATCCAAATGTTGAATTGAAGACTTTGAGTCTCACTTACCCGTTGGTACCCC 700
Db 83764 GCAACCAATCCAAATGTTGAATTGAAGACTTTGAGTCTCACTTACCCGTTGGTACCCC 83823
QY 701 CACTCTCAAGACCTCTCGAATATTCCTGTTTACCCCAAGATTTTACAGTACATC 760
Db 83824 CACTCTCAAGACCTCTCGAATATTCCTGTTTACCCCAAGATTTTACAGTACATC 83883
QY 761 TGACGAGTCTCTTGGCCAGG 781
Db 83884 TGACGAGTCTCTTGGCCAGG 83904
```

Search completed: August 27, 2005, 09:39:00  
Job time : 6002.62 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 13:32:20 ; Search time 732.44 Seconds  
(without alignments)  
16924.161 Million cell updates/sec

Title: US-09-371-347A-45

Perfect score: 2094

Sequence: 1 atgagagaggttcctgact.....ttcagatatctgtcctaa 2094

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

1: geneseqn19808:\*\n2: geneseqn19908:\*\n3: geneseqn20008:\*\n4: geneseqn20018:\*\n5: geneseqn20018:\*\n6: geneseqn20028:\*\n7: geneseqn20028:\*\n8: geneseqn20038:\*\n9: geneseqn20038:\*\n10: geneseqn20038:\*\n11: geneseqn20038:\*\n12: geneseqn20048:\*\n13: geneseqn20048:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match Length	ID	Description
1	2091	99.9	2091 11	ADM43216 Human met
2	1854	88.5	3259 5	AA65070 DNA encod
3	1854	88.5	3259 5	AA65070 DNA encod
4	1854	88.5	3259 11	ADM43206 Human ful
5	1851	88.4	2094 11	ADM43208 Human wll
6	1800	86.0	2094 11	ADM43212 Human met
7	1800	86.0	2094 11	ADM43209 Human met
8	1752	83.7	3259 13	AA58935 DNA encod
9	1701	81.2	3270 13	ADO87538 Human tum
10	1674	79.9	2091 11	ADM43214 Human met
11	1595	76.2	3256 3	AA58977 A human m
12	1544	73.7	3255 3	AA58976 A human m
13	956	45.7	3189 13	ACN42470 Human dia
14	896	42.8	3256 13	ADO39029 Human SNP
15	896	42.8	3274 13	ADO39030 Human SNP
16	562	26.8	1986 4	AA541064 cDNA enco
17	386	18.4	390 2	AA51820 Human sec
18	330	15.7	591 12	ACH73174 Human gen
19	328	15.7	379 12	ACH86905 Human gen
20	279	13.3	591 12	ACH868540 Human gen

21	277	13.2	379	12	ACH82240	ACH82240 Human gen
22	268	12.8	1663	4	AA541602	AA541602 cDNA enco
23	225	10.7	503	5	AA65069	AA65069 DNA encod
24	188	9.0	525	12	ACH67438	ACH67438 Human gen
25	175	8.4	175	12	ACH81143	ACH81143 Human gen
26	158	7.5	2475	6	AD32365	AD32365 Human lun
27	158	7.5	2475	13	AD161720	AD161720 Human gen
28	137	6.5	525	12	ACH73117	ACH73117 Human gen
29	124	5.9	175	12	ACH86848	ACH86848 Human gen
30	78	3.7	244	3	AA242736	AA242736 Human 5'
31	60	2.9	60	6	AB36264	AB36264 Human bpl
32	51	2.4	51	4	AA178548	AA178548 Human bll
33	38	1.8	1835	5	AA65071	AA65071 DNA encod
34	30	1.4	1681	11	AD131127	AD131127 Human cDN
35	26	1.2	26	3	AA58955	AA58955 PCR prime
36	26	1.2	26	3	AA58939	AA58939 PCR prime
37	26	1.2	26	6	ABX09549	ABX09549 Arteriosc
38	26	1.2	26	6	AA143713	AA143713 Pregestat
39	26	1.2	26	11	ADM43205	ADM43205 Human met
40	26	1.2	26	11	ADM43189	ADM43189 Human met
41	25	1.2	25	3	AA58952	AA58952 PCR prime
42	25	1.2	25	3	AA58937	AA58937 PCR prime
43	25	1.2	25	3	AA58947	AA58947 PCR prime
44	25	1.2	25	11	ADM43187	ADM43187 Human met
45	25	1.2	25	11	ADM43202	ADM43202 Human met

## ALIGNMENTS

RESULT 1	ADM43216	standard; cDNA; 2091 BP.
ID	ADM43216	standard; cDNA; 2091 BP.
XX	ADM43216;	
AC	ADM43216;	
XX	03-JUN-2004	(first entry)
DT	03-JUN-2004	(first entry)
XX	Human methionine synthase reductase CDS del 1726-1728 variant.	
DE	Human; ss; Methionine synthase reductase polypeptide; HsMTRR; cancer;	
XX	Human; ss; Methionine synthase reductase polypeptide; HsMTRR; cancer;	
KW	cardiovascular disease; neural tube defect; hyperhomocysteinemia;	
KM	chromosome 5p15.2-p15.3; SNP; single nucleotide polymorphism.	
XX	Homo sapiens.	
OS	Homo sapiens.	
XX	Key	Location/Qualifiers
XX	CDS	1..2091
FT		/*tag= a
FT		/product= "HsMTRRdelR559"
FT		/partial
FT		/note= "No stop codon shown"
FT	variation	replace(66, A)
FT	variation	/*tag= b
FT	variation	/standard name= "Single_nucleotide polymorphism"
FT	variation	replace(110, A)
FT	variation	/*tag= c
FT	variation	/standard_name= "Single_nucleotide polymorphism"
FT	variation	replace(1726, TTGT)
FT	variation	/*tag= d
PN	US2003082676-A1.	
XX	01-MAY-2003.	
PD	01-MAY-2003.	
XX	10-AUG-1999;	99US-00371347.
PF	10-AUG-1999;	99US-00371347.
XX	16-JAN-1998;	98US-0071622P.
XX	16-JAN-1998;	98US-0071622P.
PR	15-JAN-1999;	99US-00232028.
XX	15-JAN-1999;	99US-00232028.
XX	(GRAV/) GRAVEL R A.	
PA	(ROZE/) ROZEN R.	
PA	(LECL/) LECLERC D.	

PA (WIS/) WILSON A.  
PA (ROSE/) ROSENBLATT D.  
PI Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;  
XX WPI; 2003-576610/54.  
XX P-PSDB; ADM43217.  
XX  
XX New substantially pure nucleic acid encoding a mammalian methionine  
PT synthetase polypeptide, useful for diagnosing, preventing or  
PT treating conditions associated with altered methionine synthase activity,  
PT e.g. cancer.  
PS Disclosure; SEQ ID NO 45; 26bp; English.  
XX  
XX The invention relates to a substantially pure nucleic acid that encodes a  
CC mammalian methionine synthase reductase polypeptide, HmMTR, or that  
CC hybridizes at high stringency to a nucleic acid appearing as ADM43208 or  
CC ADM43209. Also included are a non-human animal where one or both genetic  
CC alleles encoding the methionine synthase reductase polypeptide are  
CC mutated, an antibody that specifically binds the above methionine  
CC synthase reductase polypeptide, a method of detecting the presence of the  
CC above polypeptide, a method for detecting sequence variants for  
CC methionine synthase reductase in a mammal, methods of treating or  
CC preventing cancer (or cardiovascular disease or neural tube defects) in a  
CC subject, methods of screening for a compound that modulates methionine  
CC synthase reductase biological activity and a method for detecting an  
CC increased risk of developing a neural tube defect in a mammalian embryo  
CC or foetus. The nucleic acid is useful in diagnosing, preventing or  
CC treating conditions associated with altered methionine synthase activity,  
CC such as cancer, cardiovascular disease or neural tube defects, or in  
CC screening for a compound that modulates methionine synthase reductase  
CC biological activity. Naturally occurring variants of the polypeptide are  
CC also associated with hyperhomocysteinaemia. The gene for HmMTR is  
CC located on chromosome 5p15.2-p15.3. The present sequence is the coding  
CC sequence of a variant human hsmtrr cDNA.  
XX  
XX Sequence 2091 BP; 591 A; 489 C; 480 G; 531 T; 0 U; 0 Other;  
SQ  
Query Match 99.9%; Score 2091; DB 11; Length 2091;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2091; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 421 CCGTGATGCTGGACTCTGGCCAGCCCTCAGAAAGCATTTTAAAGTCAAGCAGAGCA 480  
QY 481 GAGAGATAAGTGGCCACTCCCGGTGATCATCTGCATCTTGAAGACAGACTTTGG 540  
DB 481 GAGAGATAAGTGGCCACTCCCGGTGATCATCTGCATCTTGAAGACAGACTTTGG 540  
QY 541 AAGTCAGAGCTGACATTTGAATTCATAGTCAGAGCTTCTGAGATTGATGATTTCA 600  
DB 541 AAGTCAGAGCTGACATTTGAATTCATAGTCAGAGCTTCTGAGATTGATGATTTCA 600  
QY 601 AAGTCAGAGCTGACATTTGAATTCATAGTCAGAGCTTCTGAGATTGATGATTTCA 600  
DB 601 AAGTCAGAGCTGACATTTGAATTCATAGTCAGAGCTTCTGAGATTGATGATTTCA 600  
QY 661 ATTGAAGCTTTGAGTCTCACTTACCCGTTGATACCCCACTCTCAGAGCTTCTG 720  
DB 661 ATTGAAGCTTTGAGTCTCACTTACCCGTTGATACCCCACTCTCAGAGCTTCTG 720  
QY 721 AATATTCCTGTTTACCCCAAGATTTTACAGATCATCTGAGAGATCTTGGCCAG 780  
DB 721 AATATTCCTGTTTACCCCAAGATTTTACAGATCATCTGAGAGATCTTGGCCAG 780  
QY 781 GAGGAAGCCAGATCTGATCTGAGAGATCCAGTTTTCAGTGGCAATTTGAAAG 840  
DB 781 GAGGAAGCCAGATCTGATCTGAGAGATCCAGTTTTCAGTGGCAATTTGAAAG 840  
QY 841 GCAGTTCACTTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
DB 841 GCAGTTCACTTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
QY 901 TCAATATCAGACTTTTCTATCATGCTGAGATGCTTCAAGCTGATCTGCTTAA 960  
DB 901 TCAATATCAGACTTTTCTATCATGCTGAGATGCTTCAAGCTGATCTGCTTAA 960  
QY 961 GATTCGAGATCAAGGCTTCTCAAGATGCTGAGATGCTGAGATGCTGAGATGCTG 1020  
DB 961 GATTCGAGATCAAGGCTTCTCAAGATGCTGAGATGCTGAGATGCTGAGATGCTG 1020  
QY 1021 GTCTTTTGAATAAATGAAGCAGACACAAAGAAAGAGATCTTACCCAGATATA 1080  
DB 1021 GTCTTTTGAATAAATGAAGCAGACACAAAGAAAGAGATCTTACCCAGATATA 1080  
QY 1081 CCGCGGATGTTCTCTCAAGTCAATTTTCTGCTGCTTGAATTCGAGATTTCT 1140  
DB 1081 CCGCGGATGTTCTCTCAAGTCAATTTTCTGCTGCTTGAATTCGAGATTTCT 1140  
QY 1141 AAAAAGCATTTTGGAGCCCTTGGAGCTATACAGATGACAGTCTGAAAAGCCAG 1200  
DB 1141 AAAAAGCATTTTGGAGCCCTTGGAGCTATACAGATGACAGTCTGAAAAGCCAG 1200  
QY 1201 CTACAGAGCTGTCAGTAAACAAAGGAGGAGCCGATTAAGCCGTTTGAAGATGCC 1260  
DB 1201 CTACAGAGCTGTCAGTAAACAAAGGAGGAGCCGATTAAGCCGTTTGAAGATGCC 1260  
QY 1261 TGTGCTGCTTGTGATCTCTCTGCTGCTTCTTCTGAGCAGCAGCAGTCTC 1320  
DB 1261 TGTGCTGCTTGTGATCTCTCTGCTGCTTCTTCTGAGCAGCAGCAGTCTC 1320  
QY 1321 CTGCTGGAATCTTCTTAACTTCAACCAAGCAGATATTCGAGAGCTCAAGTTA 1380  
DB 1321 CTGCTGGAATCTTCTTAACTTCAACCAAGCAGATATTCGAGAGCTCAAGTTA 1380  
QY 1381 TTTCAACCAAGAGCTCAATTTTGTCTTCAACATTTGGAATTTGTCTACTGAC 1440  
DB 1381 TTTCAACCAAGAGCTCAATTTTGTCTTCAACATTTGGAATTTGTCTACTGAC 1440  
QY 1441 ACAAGATCTGCGGAGAGGAGATGATGATGATGATGATGATGATGATGATGAT 1500  
DB 1441 ACAAGATCTGCGGAGAGGAGATGATGATGATGATGATGATGATGATGATGAT 1500  
QY 1501 CTTGAGCAGAAACATATGATGATGATGATGATGATGATGATGATGATGATGAT 1560  
DB 1501 CTTGAGCAGAAACATATGATGATGATGATGATGATGATGATGATGATGATGAT 1560

QY 1561 TCCATCTCTCTGGAACAACAATTCTTTCACCTTACGAGATGACCCCTCAATCCCCATC 1620  
DB 1561 TCCATCTCTCTGGAACAACAATTCTTTCACCTTACGAGATGACCCCTCAATCCCCATC 1620  
QY 1621 ATAAATGTTGGGTCAGGAACCGGCAATGCGGCTTATTTGGGTTCTTACACATAGAGAG 1680  
DB 1621 ATAAATGTTGGGTCAGGAACCGGCAATGCGGCTTATTTGGGTTCTTACACATAGAGAG 1680  
QY 1681 AAATCTCAAGAACACACACCCAGATGGAATTTTGGAGCAATGTTGGTCTGAGG 1740  
DB 1681 AAATCTCAAGAACACACACCCAGATGGAATTTTGGAGCAATGTTGGTCTGAGG 1740  
QY 1741 CATTAGGATGAGGATTTATCTATTCAGAAAAGCTCAGACATTTCTTAAAGCATGGGATC 1800  
DB 1741 CATTAGGATGAGGATTTATCTATTCAGAAAAGCTCAGACATTTCTTAAAGCATGGGATC 1800  
QY 1801 TTTAACTCATTTAAAGGTTCTCTTCAAGAGATGCTCTGTTGGGAGAGAGAGCCCA 1860  
DB 1801 TTTAACTCATTTAAAGGTTCTCTTCAAGAGATGCTCTGTTGGGAGAGAGAGCCCA 1860  
QY 1861 GCAAAATATGTACAGACACATCCAGCTTCATGSCCAGAGGTGGCGAGATCTCTC 1920  
DB 1861 GCAAAATATGTACAGACACATCCAGCTTCATGSCCAGAGGTGGCGAGATCTCTC 1920  
QY 1921 CAGGAGAACGGCCATTTATGTTGTGAGAGATGCAAAAGATTTGGCCAAAGATGTACAT 1980  
DB 1921 CAGGAGAACGGCCATTTATGTTGTGAGAGATGCAAAAGATTTGGCCAAAGATGTACAT 1980  
QY 1981 GATGCGCTTGTGCAATTAATTAAGCAAGAGGTTGAGTTGAAAACTAGAGCATGAAA 2040  
DB 1981 GATGCGCTTGTGCAATTAATTAAGCAAGAGGTTGAGTTGAAAACTAGAGCATGAAA 2040  
QY 2041 ACCCTGGCCACTTTAAAGAGAAAAACGCTACCTTCAGATATTTGTCA 2091  
DB 2041 ACCCTGGCCACTTTAAAGAGAAAAACGCTACCTTCAGATATTTGTCA 2091

## RESULT 2

AA565070  
ID AA565070 standard; cDNA; 3259 BP.

AA565070;

AC 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #874.

KM Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

PN MO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001MO-US008631.

PR 31-MAR-2000; 2000US-00540217.  
PR 23-AUG-2000; 2000US-00649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

P-PSDB; ABG00883.

PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.

XX  
PS Claim 1, SEQ ID NO 874; 103pp; English.  
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostic as expressed sequence tags for identifying expressed  
CC genes. (II) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AA564197-AA594564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX

Sequence 3259 BP; 944 A; 695 C; 674 G; 946 T; 0 U; 0 Other;

Query Match 88.5%; Score 1854; DB 5; Length 3259;

Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2094; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 ATGAGAGGTTCTGTTACTATATGCTACACAGCAGGAGCAGGCAAAAGCCATCGCAGAA 60  
DB 80 ATGAGAGGTTCTGTTACTATATGCTACACAGCAGGAGCAGGCAAAAGCCATCGCAGAA 139  
QY 61 GAAATGTGTACAGACGCTGTGTACATGATTTTCTGAGATCTTCACTGATTAATGAA 120  
DB 140 GAAATGTGTACAGACGCTGTGTACATGATTTTCTGAGATCTTCACTGATTAATGAA 199  
QY 121 TCCGATTAATGACCTTAATAAACCGAACAAGCTCTCTGTTGTTGTTCTTACACAG 180  
DB 200 TCCGATTAATGACCTTAATAAACCGAACAAGCTCTCTGTTGTTGTTCTTACACAG 259  
QY 181 GGCACCGGAGACCCACCGCAGACAGCCCGCAGTTGTTAAGGAATACAGAACCA 240  
DB 260 GGCACCGGAGACCCACCGCAGACAGCCCGCAGTTGTTAAGGAATACAGAACCA 319  
QY 241 CTGCGGTTGATTTCTTGTCTCACTGCGGTATGAGTTACTGGGTTCTCGGTATTCAGAA 300  
DB 320 CTGCGGTTGATTTCTTGTCTCACTGCGGTATGAGTTACTGGGTTCTCGGTATTCAGAA 379  
QY 301 TACACCTACTTTTGCATGAGGGGGGAAATTAATTAACGACTTCAAGAGCTTGAAGCC 360  
DB 380 TACACCTACTTTTGCATGAGGGGGGAAATTAATTAACGACTTCAAGAGCTTGAAGCC 439  
QY 361 CGGCATTTCTATGACACTGAGACATGAGATGATGTTGAGTTTAAAGCTTGGTTGAG 420  
DB 440 CGGCATTTCTATGACACTGAGACATGAGATGATGTTGAGTTTAAAGCTTGGTTGAG 499  
QY 421 CCGTGAATGCTGGACTCTGCGCAGAGCCCTCAGAAACATTTTAAAGTCAAGAGAGCA 480  
DB 500 CCGTGAATGCTGGACTCTGCGCAGAGCCCTCAGAAACATTTTAAAGTCAAGAGAGCA 559  
QY 481 GAGAGATTAAGTGGGCACTCCCGGTGAGATCACTGCAATCTTGAAGACAGACTTTG 540  
DB 560 GAGAGATTAAGTGGGCACTCCCGGTGAGATCACTGCAATCTTGAAGACAGACTTTG 619  
QY 541 AAGTCAGAGTGTACATTAATCAATCAAGTCAAGTCTTGAAGTTGATTCAGGATTCAGGA 600  
DB 620 AAGTCAGAGTGTACATTAATCAATCAAGTCAAGTCTTGAAGTTGATTCAGGATTCAGGA 679  
QY 601 AGAAAGATTTCTGAGGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATCCATGTTGTA 660

Db 680 AGAAGATTCTGAGTTTGAAGCAAAATGCACTGAACGCAACCAATCCAAATGTTGTA 739  
Qy 661 ATTGAAGACTTTGAGTCTCTCACTTACCCCTGCGTAACCCCACTCTCAAGAGCTCTG 720  
Db 740 ATTGAAGACTTTGAGTCTCTCACTTACCCCTGCGTAACCCCACTCTCAAGAGCTCTG 739  
Qy 721 AATATTCCTGGTTTACCCCAAAATATTTACAGGTACATCTGACGAGTCTTTGGCCAG 780  
Db 800 AATATTCCTGGTTTACCCCAAAATATTTACAGGTACATCTGACGAGTCTTTGGCCAG 859  
Qy 781 GAGAAAGCCAAAGATCTGTAAGTCTGAGAGTCCAGTCTTTTCAAGTCCAAATTTCAAG 840  
Db 860 GAGAAAGCCAAAGATCTGTAAGTCTGAGAGTCCAGTCTTTTCAAGTCCAAATTTCAAG 919  
Qy 841 GCAGTTCAACTTAATAAGAAAGTGAATGCAATGCACTGCTGTAGTAATTTGACAT 900  
Db 920 GCAGTTCAACTTAATAAGAAAGTGAATGCAATGCACTGCTGTAGTAATTTGACAT 979  
Qy 901 TCAAAATACAGACTTTTCTTATCAAGCTGAGAGTCTTCAAGCTGATCTGCTTAACAT 960  
Db 980 TCAAAATACAGACTTTTCTTATCAAGCTGAGAGTCTTCAAGCTGATCTGCTTAACAT 1039  
Qy 961 GATTCTGAGGTACAAAGCTTCTCAAGAGCTGAGAGTCTTCAAGTAAAGAGAGCACTGC 1020  
Db 1040 GATTCTGAGGTACAAAGCTTCTCAAGAGCTGAGAGTCTTCAAGTAAAGAGAGCACTGC 1099  
Qy 1021 GTCTTTTGAATAAAGAGAGACACAAAGAAAGAGACTTATCCCAAGCATATA 1080  
Db 1100 GTCTTTTGAATAAAGAGAGACACAAAGAAAGAGACTTATCCCAAGCATATA 1159  
Qy 1081 CTTGCGGAGATGTTCTCTCAAGTCTTATCTGAGTCTTGAATCCGAGCAATTCCT 1140  
Db 1160 CTTGCGGAGATGTTCTCTCAAGTCTTATCTGAGTCTTGAATCCGAGCAATTCCT 1219  
Qy 1141 AAAAAGCATTTTTCGAGGCTTGTGACATATACAGTACAGTGTGAAAAAGCCGAG 1200  
Db 1220 AAAAAGCATTTTTCGAGGCTTGTGACATATACAGTACAGTGTGAAAAAGCCGAG 1279  
Qy 1201 CTAAGAGAGTGTGACAGTAAACAAGGGGAGCCGATTAACCGCTTTGTACAGATGCC 1260  
Db 1280 CTAAGAGAGTGTGACAGTAAACAAGGGGAGCCGATTAACCGCTTTGTACAGATGCC 1339  
Qy 1261 TGTGCTGCTTTGTGATCTCTCTCTGCTTCCCTTCTTGCCAGCCACATCAAGTCTC 1320  
Db 1340 TGTGCTGCTTTGTGATCTCTCTCTGCTTCCCTTCTTGCCAGCCACATCAAGTCTC 1399  
Qy 1321 CTGCTGGAACATCTTCTTAACTCAACCCAGACATATTCGTGCAAGCTCAAGTTTA 1380  
Db 1400 CTGCTGGAACATCTTCTTAACTCAACCCAGACATATTCGTGCAAGCTCAAGTTTA 1459  
Qy 1381 TTTCAACCAGAAAGCTCAATTTGTCTTCAACATTTGTGAAATTTCTGTACTGCCACA 1440  
Db 1460 TTTCAACCAGAAAGCTCAATTTGTCTTCAACATTTGTGAAATTTCTGTACTGCCACA 1519  
Qy 1441 ACAGAGTTTCTGCGGAGGAGATATTAACAGCTGCTGCTTGTGTTGCTTCAAGT 1500  
Db 1520 ACAGAGTTTCTGCGGAGGAGATATTAACAGCTGCTGCTTGTGTTGCTTCAAGT 1579  
Qy 1501 CTTCAAGCAAAATACATGATCCATGAAAGACAGCGGAAAGCCCTGCTCTTAAGATA 1560  
Db 1580 CTTCAAGCAAAATACATGATCCATGAAAGACAGCGGAAAGCCCTGCTCTTAAGATA 1639  
Qy 1561 TCCATCTCTCTGGAACAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
Db 1640 TCCATCTCTCTGGAACAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1699  
Qy 1621 ATATATGTTGGTCCAGAGACCGGATATGCCCCGTTTATTTGGTTCTTACAAATAGAG 1680  
Db 1700 ATATATGTTGGTCCAGAGACCGGATATGCCCCGTTTATTTGGTTCTTACAAATAGAG 1759  
Qy 1681 AAATCTCAAGAAACAACCCAGATGAAATTTTGGACCAATGTS--GTTTTTGGCTGC 1737  
Db 1760 AAATCTCAAGAAACAACCCAGATGAAATTTTGGACCAATGTS--GTTTTTGGCTGC 1819

Qy 1738 AGCATTAAGATTAAGGATTATCTATTACAGAAAAAGCTCAGACATTTCTTAAAGATGG 1797  
Db 1820 AGCATTAAGATTAAGGATTATCTATTACAGAAAAAGCTCAGACATTTCTTAAAGATGG 1879  
Qy 1798 ATCTTAATCACTTAAGATTTCTTCTCAAGAGATGCTCTGTTGGAGAGAGGAAGCC 1857  
Db 1880 ATCTTAATCACTTAAGATTTCTTCTCAAGAGATGCTCTGTTGGAGAGAGGAAGCC 1939  
Qy 1858 CCAGCAAGATTAATGACAAACAACATCCAGCTTCAAGGACAGAGGTGCGAATCTCTC 1917  
Db 1940 CCAGCAAGATTAATGACAAACAACATCCAGCTTCAAGGACAGAGGTGCGAATCTCTC 1999  
Qy 1918 CTTCAAGAGAACGCGCATATTTATGTGTGAGATGCAAGAAATATGCGCAAGATGTA 1977  
Db 2000 CTTCAAGAGAACGCGCATATTTATGTGTGAGATGCAAGAAATATGCGCAAGATGTA 2059  
Qy 1978 CATGATGCCCCCTTGTGCAATTAATTAACCAAGAGCTTGGAGTTGAAATCATGAAGCAATG 2037  
Db 2060 CATGATGCCCCCTTGTGCAATTAATTAACCAAGAGCTTGGAGTTGAAATCATGAAGCAATG 2119  
Qy 2038 AAAACCTGGCCACTTAAAGAAAGAAACGCTACCTTCAAGATATTTGTCATTA 2094  
Db 2120 AAAACCTGGCCACTTAAAGAAAGAAACGCTACCTTCAAGATATTTGTCATTA 2176

RESULT 3  
AAC91226  
ID AAC91226 standard; DNA; 3259 BP.  
XX  
AC AAC91226;  
XX  
DT 20-MAR-2001 (first entry)  
XX  
DE Human schizophrenia related gene SEQ ID NO: 23.  
XX  
KW Human; schizophrenia; developmental disorder; spina bifida cystica;  
KW Tourette's syndrome; bipolar illness; autism; conduct disorder;  
KW attention deficit disorder; obsessive compulsive disorder;  
KW chronic multiple tic syndrome; learning disorder; polymorphism; ds.  
XX  
OS Homo sapiens.  
XX  
PN MO200071754-A1.  
XX  
PD 30-NOV-2000.  
XX  
PF 24-MAY-2000; 2000NC-US014354.  
XX  
PR 25-MAY-1999; 99US-00318448.  
XX  
PA (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.  
XX  
PI Johnson WG, Stenroos ES;  
XX  
DR WPI; 2001-025174/03.  
XX  
PT Diagnosing a developmental disorder, e.g. schizophrenia, by forming  
PT datasets (DS) of genetic (e.g. genotypes of folate metabolism alleles)  
PT and environmental variables affecting an individual and then comparing  
PT these DS with reference DS.  
XX  
PS Disclosure; Page 142-143, 156pp; English.  
XX  
CC The present invention provides a novel method of estimating the  
CC susceptibility of an individual to a developmental disorder using genetic  
CC and environmental variables. The method can be used in the diagnosis,  
CC prevention and treatment of disorders such as schizophrenia, spina bifida  
CC cystica, Tourette's syndrome, bipolar illness, autism, conduct disorders,  
CC attention deficit disorder, obsessive compulsive disorder, chronic  
CC multiple tic syndrome and learning disorders such as dyslexia  
XX  
SQ Sequence 3259 BP; 944 A; 695 C; 674 G; 946 T; 0 U; 0 Other;

Query Match 88.5%; Score 1854; DB 5; Length 3259;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 2094; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

1 ATGAGAGGTTCTGTACTATATGCTACACAGCAGGAGCAGGCAAAAGCCATCGCAGA 60  
 80 ATGAGAGGTTCTGTACTATATGCTACACAGCAGGAGCAGGCAAAAGCCATCGCAGA 139  
 61 GAAATGTGTACCAAGCTGTGTACATGTGATTTCTGCGATCTTCACTGTATATGTGA 120  
 140 GAAATGTGTACCAAGCTGTGTACATGTGATTTCTGCGATCTTCACTGTATATGTGA 199  
 121 TCCGATTAAGTATGACCTTAATAAACCGAAGAGCTCTCTGTGTGTGTGTCTACACAG 180  
 200 TCCGATTAAGTATGACCTTAATAAACCGAAGAGCTCTCTGTGTGTGTGTCTACACAG 259  
 181 GGCACCGGAGACCCACCGACACAGCCCGCAAGTTGTTAAGGAAATACAGAACCAACA 240  
 260 GGCACCGGAGACCCACCGACACAGCCCGCAAGTTGTTAAGGAAATACAGAACCAACA 319  
 241 CTGCGGTGATTTCTTGTCTCACTGCGGTATGGGTACTGGGTCTCGGTATTCAGAA 300  
 320 CTGCGGTGATTTCTTGTCTCACTGCGGTATGGGTACTGGGTCTCGGTATTCAGAA 379  
 301 TACACTACTTTTGGCAATGGGGGAGATTAATTGATTAACGACTTCAAGAGCTTGAGGC 360  
 380 TACACTACTTTTGGCAATGGGGGAGATTAATTGATTAACGACTTCAAGAGCTTGAGGC 439  
 361 CGGCAATTTCTATGACACTGGAACATGACATGACTGTAGTGTAGAACCTTGTGTAG 420  
 440 CGGCAATTTCTATGACACTGGAACATGACATGACTGTAGTGTAGAACCTTGTGTAG 499  
 421 CCGTGATTTGCTGGAATCTGGCCAGCCCTCAGAAAGCATTTTAGTCAAGCAGAGACA 480  
 500 CCGTGATTTGCTGGAATCTGGCCAGCCCTCAGAAAGCATTTTAGTCAAGCAGAGACA 559  
 481 GAGGAGATTAAGTGGGCACTCCCGGAGATCACTCGCATCCCTGAGGACAGACCTTGTG 540  
 560 GAGGAGATTAAGTGGGCACTCCCGGAGATCACTCGCATCCCTGAGGACAGACCTTGTG 619  
 541 AAGTCAGAGCTGTACACATTTGATCTCAAGTCGAGCTTCTAGATTCGATGATTCAGGA 600  
 620 AAGTCAGAGCTGTACACATTTGATCTCAAGTCGAGCTTCTAGATTCGATGATTCAGGA 679  
 601 AAGGAGATTTGAGGTTTGAAGCAAAATGCAAGTAAACAGCAACCAATCCAAATGTTGA 660  
 680 AAGGAGATTTGAGGTTTGAAGCAAAATGCAAGTAAACAGCAACCAATCCAAATGTTGA 739  
 661 ATTGAAGACTTTGAGTCTCTCACTTACCCGTTGGTACCCCACTCTCAAGAGCTCTCTG 720  
 740 ATTGAAGACTTTGAGTCTCTCACTTACCCGTTGGTACCCCACTCTCTCAAGAGCTCTCTG 799  
 721 AATATTCCTGTGTTACCCCAAGATATTTACAGGTACATCTGAGAGAGTCTTGTGCGAG 780  
 800 AATATTCCTGTGTTACCCCAAGATATTTACAGGTACATCTGAGAGAGTCTTGTGCGAG 859  
 781 GAGGAAAGCCAGATCTGTGACTTCAAGAGATCAAGTTTTCAGTCCCAATTTCAAG 840  
 860 GAGGAAAGCCAGATCTGTGACTTCAAGAGATCAAGTTTTCAGTCCCAATTTCAAG 919  
 841 GAGGTTCACTTCTACAGATGATGCTTAATAAACCACTCTGTGTGTATGATTCAGATTC 900  
 920 GAGGTTCACTTCTACAGATGATGCTTAATAAACCACTCTGTGTGTATGATTCAGATTC 979  
 901 TCAATATCAGACTTTCTATCAGCTGAGATGCTTCAAGCTGTGATTCGCTTACAGT 960  
 980 TCAATATCAGACTTTCTATCAGCTGAGATGCTTCAAGCTGTGATTCGCTTACAGT 1039  
 961 GATTCTGAGTACAAAGCTTCTCAAGACCTGAGAGTGAAGATTAAGAGAGACCTGC 1020  
 1040 GATTCTGAGTACAAAGCTTCTCAAGACCTGAGAGTGAAGATTAAGAGAGACCTGC 1099

1021 GTCTTTGAAAATAAAGCAGACACAAAGAAAGAGCTACTTACCAGCATATTA 1080  
 1100 GTCTTTGAAAATAAAGCAGACACAAAGAAAGAGCTACTTACCAGCATATTA 1159  
 1081 CTTGGGGATGTTCTCTCAAGTTCAATTTTACCTGTGTCTTGAATCCGACCAATTCCT 1140  
 1160 CTTGGGGATGTTCTCTCAAGTTCAATTTTACCTGTGTCTTGAATCCGACCAATTCCT 1219  
 1141 AAAAAGCATTTTTCGAGCCCTTGTGACATTAACAGTACAGTGTCTGAAAAGCAGAG 1200  
 1220 AAAAAGCATTTTTCGAGCCCTTGTGACATTAACAGTACAGTGTCTGAAAAGCAGAG 1279  
 1201 CTACAGAGCTGTGACATTAACAGAGGAGCCGATTAATAGCCGCTTGTACAGATGCC 1260  
 1280 CTACAGAGCTGTGACATTAACAGAGGAGCCGATTAATAGCCGCTTGTACAGATGCC 1339  
 1261 TGTGCTGCTGTTGTGATCTCTCTCGCTTTCCTTCTTGGCAGCCACACTCACTCTC 1320  
 1340 TGTGCTGCTGTTGTGATCTCTCTCGCTTTCCTTCTTGGCAGCCACACTCACTCTC 1399  
 1321 CTGCTCGAATCATCTTCTTAACTTCAACCCAGACATTAATGTTGAGAGCTCAAGTTTA 1380  
 1400 CTGCTCGAATCATCTTCTTAACTTCAACCCAGACATTAATGTTGAGAGCTCAAGTTTA 1459  
 1381 TTTCAACCCAGAAGCTCCATTTTGTCTCAACATTTGAGAAATTTCTGTACTGACACA 1440  
 1460 TTTCAACCCAGAAGCTCCATTTTGTCTTCAACATTTGAGAAATTTCTGTACTGACACA 1519  
 1441 ACGAGTTCTGCGAAGGAGATGATGACAGCTGTGCTGTGTGTGTGCTTCAAGTT 1500  
 1520 ACGAGTTCTGCGAAGGAGATGATGACAGCTGTGCTGTGTGTGTGCTTCAAGTT 1579  
 1501 CTTGAGCCAAATATGATGATCCCATGAGACAGGAGGAGAAAGCCCTGCTCTTAAGATA 1560  
 1580 CTTGAGCCAAATATGATGATCCCATGAGACAGGAGGAGAAAGCCCTGCTCTTAAGATA 1639  
 1561 TCCATCTCTCTCGAACAACAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
 1640 TCCATCTCTCTCGAACAACAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1699  
 1621 ATTAATGTGGTTCAGAGAACCGGCAATAGCCCGTTTATGGTTCCTTACACATAGAGAG 1680  
 1700 ATTAATGTGGTTCAGAGAACCGGCAATAGCCCGTTTATGGTTCCTTACACATAGAGAG 1759  
 1681 AAATCTCAAGAACCAACCCAGATGAGAAATTTTGAAGCAATG---GTTTTTGGCTGC 1737  
 1760 AAATCTCAAGAACCAACCCAGATGAGAAATTTTGAAGCAATG---GTTTTTGGCTGC 1819  
 1738 AAGCATTAAGGATTAAGGATTAATCTTATCAGAAAGAGCTCAGACATTTCTTAAAGATGG 1797  
 1820 AAGCATTAAGGATTAAGGATTAATCTTATCAGAAAGAGCTCAGACATTTCTTAAAGATGG 1879  
 1798 ATCTTAATCTCATTAAGGTTTCTTCTCAGAGATGCTCTGTGGGAGAGAGAGCC 1857  
 1880 ATCTTAATCTCATTAAGGTTTCTTCTCAGAGATGCTCTGTGGGAGAGAGAGCC 1939  
 1858 CAGCAAAATATGATCAAGAACCAATCCAGCTTCAATGAGCAGAGTGGCAGAAATCTCTC 1917  
 1940 CAGCAAAATATGATCAAGAACCAATCCAGCTTCAATGAGCAGAGTGGCAGAAATCTCTC 1999  
 1918 CTTCAGAGAGAGGAGCAATTAATGATGATGAGATGCAAGAAATATGAGCAAGATGTA 1977  
 2000 CTTCAGAGAGAGGAGCAATTAATGATGATGAGATGCAAGAAATATGAGCAAGATGTA 2059  
 1978 CATGATGCCCTTGTGCAATTAATGCAAGAGGTTGAGGTTGAAAATCTAAGCAATG 2037  
 2060 CATGATGCCCTTGTGCAATTAATGCAAGAGGTTGAGGTTGAAAATCTAAGCAATG 2119  
 2038 AAAACCTGAGCACTTTAAAGAAAGAAAGCCCTCAAGATTTTGTGCTATA 2094  
 2120 AAAACCTGAGCACTTTAAAGAAAGAAAGCCCTCAAGATTTTGTGCTATA 2176

RESULT 4  
ADM43206  
ID ADM43206 standard; cDNA; 3259 BP.  
XX  
AC ADM43206;  
XX  
DT 03-JUN-2004 (first entry)  
XX  
DE Human full length cDNA encoding methionine synthase reductase.  
XX  
KW Human; ss; gene; Methionine synthase reductase polypeptide; HemTR; cancer; cardiovascular disease; neural tube defect; hyperhomocysteinemia; chromosome 5p15.2-p15.3; SNP; single nucleotide polymorphism.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 80..2176  
FT /\*tag= a  
FT /product= "hemTR"  
FT /replac= (145, A)  
FT /\*tag= b  
FT /strand name= "Single\_nucleotide polymorphism"  
FT /replac= (189, A)  
FT /\*tag= c  
FT /strand name= "Single\_nucleotide polymorphism"  
XX  
PN US2003082676-A1.  
XX  
PD 01-MAY-2003.  
XX  
XX 10-AUG-1999; 99US-00371347.  
XX  
PR 15-JAN-1998; 98US-0071622P.  
XX 15-JAN-1999; 99US-00232028.  
XX  
PA (GRAY/) GRAVEL R A.  
PA (ROZE/) ROZEN R.  
PA (LECL/) LECLERC D.  
PA (WILS/) WILSON A.  
PA (ROSE/) ROSENBLATT D.  
XX  
PI Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;  
XX  
XX MPI; 2003-57610/54.  
XX P-PSDB; ADM43207.  
XX  
PT New substantially pure nucleic acid encoding a mammalian methionine  
PT synthase reductase polypeptide, useful for diagnosing, preventing or  
PT treating conditions associated with altered methionine synthase activity,  
PT e.g. cancer.  
XX  
XX Example 2; SEQ ID NO 24; 26pp; English.  
XX  
XX The invention relates to a substantially pure nucleic acid that encodes a  
XX mammalian methionine synthase reductase polypeptide, HemTR, or that  
XX hybridizes at high stringency to a nucleic acid appearing as ADM43208 or  
XX ADM43209. Also included are a non-human animal where one or both genetic  
XX alleles encoding the methionine synthase reductase polypeptide are  
XX mutated, an antibody that specifically binds the above methionine  
XX synthase reductase polypeptide, a method of detecting the presence of the  
XX above polypeptide, a method for detecting sequence variants for  
XX methionine synthase reductase in a mammal, methods of treating or  
XX preventing cancer (or cardiovascular disease or neural tube defects) in a  
XX subject, methods of screening for a compound that modulates methionine  
XX synthase reductase biological activity and a method for detecting an  
XX increased risk of developing a neural tube defect in a mammalian embryo  
XX or foetus. The nucleic acid is useful in diagnosing, preventing or  
XX treating conditions associated with altered methionine synthase activity,  
XX such as cancer, cardiovascular disease or neural tube defects, or in  
XX screening for a compound that modulates methionine synthase reductase  
XX biological activity. Naturally occurring variants of the polypeptide are

CC also associated with hyperhomocysteinemia. The gene for HemTR is  
CC located on chromosome 5p15.2-p15.3. The present sequence is full length  
CC sequence of the wild-type human hemTR cDNA.  
XX  
SQ Sequence 3259 BP; 944 A; 695 C; 674 G; 946 T; 0 U; 0 Other;  
Query Match 88.5%; Score 1854; DB 11; Length 3259;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2094; Conservative 0; Mismatches 0; Indels 3; Gaps 1;  
QY 1 ATGAGAGGTTTCTGTATCATATGCTACACAGCAGGACGCAAGCCATCGAGAA 60  
DB 80 ATGAGAGGTTTCTGTATCATATGCTACACAGCAGGACGCAAGCCATCGAGAA 139  
QY 61 GAAATGTGAGCAAGCTGTGTACATGATTTTCTGAGATCTTCACTGATTAAGTAA 120  
DB 140 GAAATGTGAGCAAGCTGTGTACATGATTTTCTGAGATCTTCACTGATTAAGTAA 199  
QY 121 TCCGATTAAGTATGACTTAATAAACCAGAACGCTCTTGTGTGTGTTTCTACACG 180  
DB 200 TCCGATTAAGTATGACTTAATAAACCAGAACGCTCTTGTGTGTGTTTCTACACG 259  
QY 181 GGCACCGAGACCCACCCGACACAGCCGCAAGTTGTTAAGGAATACAGAACCAACA 240  
DB 260 GGCACCGAGACCCACCCGACACAGCCGCAAGTTGTTAAGGAATACAGAACCAACA 319  
QY 241 CTGCGGTTGATTTCTTGTCTCACTGCGGTATGAGTTTCTGCGTCTCGGTATTCAGAA 300  
DB 320 CTGCGGTTGATTTCTTGTCTCACTGCGGTATGAGTTTCTGCGTCTCGGTATTCAGAA 379  
QY 301 TACACTTACTTTTGCAATGGGGGGAATATTTGAATTAACGACTTAAGAGCTTGGAGCC 360  
DB 380 TACACTTACTTTTGCAATGGGGGGAATATTTGAATTAACGACTTGAAGAGCTTGGAGCC 439  
QY 361 CGGATTTCTATGACACTGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 420  
DB 440 CGGATTTCTATGACACTGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 499  
QY 421 CCGTGATTTGCTGACTGTGGCCAGCCCTCAGAACATTTTATGATCAAGAGAGACAA 480  
DB 500 CCGTGATTTGCTGACTGTGGCCAGCCCTCAGAACATTTTATGATCAAGAGAGACAA 559  
QY 481 GAGGATTAATGAGGCGCACTCCCGGTGATCATCTGATCTTGTGAGACAGACCTTGTG 540  
DB 560 GAGGATTAATGAGGCGCACTCCCGGTGATCATCTGATCTTGTGAGACAGACCTTGTG 619  
QY 541 AAGTCAGAGCTGCTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
DB 620 AAGTCAGAGCTGCTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 679  
QY 601 AGAAGGATTTGAGGTTTGAAGCAAAATGCAAGTGAAGCAAGCAAGCAAGTGTGTA 660  
DB 680 AGAAGGATTTGAGGTTTGAAGCAAAATGCAAGTGAAGCAAGCAAGCAAGTGTGTA 729  
QY 661 ATTGAAGACTTTGAGTCTTCACTTACCCGTTGGTATCCCACTTCTCAAGGCTCTGTG 720  
DB 740 ATTGAAGACTTTGAGTCTTCACTTACCCGTTGGTATCCCACTTCTCAAGGCTCTGTG 799  
QY 721 AATATTCCTGTTTACCCCAAGAAATTTTACAGGTATCTGACAGAGTCTTGGCCAG 780  
DB 800 AATATTCCTGTTTACCCCAAGAAATTTTACAGGTATCTGACAGAGTCTTGGCCAG 859  
QY 781 GAGGAAGCCAGATATCTGTGACTTGAAGATTCAGATTCAGATTCAGATTCAGATTCAG 840  
DB 860 GAGGAAGCCAGATATCTGTGACTTGAAGATTCAGATTCAGATTCAGATTCAGATTCAG 919  
QY 841 GCGATTCACCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
DB 920 GCGATTCACCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 979  
QY 901 TCAATATGACATTTTCTATGAGCTTGAAGATGCTTCAAGGATGATGCTTACCTTACAGT 960  
DB 980 TCAATATGACATTTTCTATGAGCTTGAAGATGCTTCAAGGATGATGCTTACCTTACAGT 1039



```
QY 961 GATTCTGAGGTACAAAGCCCTACTCCAAAGACTGACCTTGAAAGATTAAGAGACACTGC 1020
DB 1040 GATTCTGAGGTACAAAGCCCTACTCCAAAGACTGACCTTGAAAGATTAAGAGACACTGC 1099
QY 1021 GTCTCTTTGAAAATAAAGGACAGACAAAGAAAGAGAGCTTACCTCCAGCATATA 1080
DB 1100 GTCTCTTTGAAAATAAAGGACAGACAAAGAAAGAGAGCTTACCTCCAGCATATA 1159
QY 1081 CCTGCGGGAGTGTCTCTCCAGTTCAATTTTAACTGTGTCTTGAATTCGAGCAATTCCT 1140
DB 1160 CCTGCGGGAGTGTCTCTCCAGTTCAATTTTAACTGTGTCTTGAATTCGAGCAATTCCT 1219
QY 1141 AAAAAGGCAATTTTGGAGCCCTGTGACTATACAGTGAAGTGTGCTGAAAAAGGCAAG 1200
DB 1220 AAAAAGGCAATTTTGGAGCCCTGTGACTATACAGTGAAGTGTGCTGAAAAAGGCAAG 1279
QY 1201 CTACAGAGACTGTGACGTAAACAAGGGGACGCCGATTAATAGCCGCTTTGTACGAGATGCC 1260
DB 1280 CTACAGAGACTGTGACGTAAACAAGGGGACGCCGATTAATAGCCGCTTTGTACGAGATGCC 1339
QY 1261 TGTGCTGCTGTGTGGATCTCTCTGCTTCCCTTCTTGCCAGCCACACTCACTCTC 1320
DB 1340 TGTGCTGCTGTGTGGATCTCTCTGCTTCCCTTCTTGCCAGCCACACTCACTCTC 1399
QY 1321 CTGCTGGAACATCTTCTAACTTCAACCCAGACATTCGTGACAGCTCAAGTTTA 1459
DB 1400 CTGCTGGAACATCTTCTAACTTCAACCCAGACATTCGTGACAGCTCAAGTTTA 1459
QY 1381 TTTCACCCAGAAAGCTCCATTTTGTCTCAACATTTGTGAAATTTCTGTCTACCTGCACA 1440
DB 1460 TTTCACCCAGAAAGCTCCATTTTGTCTCAACATTTGTGAAATTTCTGTCTACCTGCACA 1519
QY 1441 ACAAGAGTTCTGGGAAAGGAGATATGATACAGGCTGGCTGCTTGTGCTTCACTT 1500
DB 1520 ACAAGAGTTCTGGGAAAGGAGATATGATACAGGCTGGCTGCTTGTGCTTCACTT 1579
QY 1501 CTTTCAAGCAAAATATATGATATCCCAATGAAAGACAGGGGAAAGCCCTGCTCTTAAGATA 1560
DB 1580 CTTTCAAGCAAAATATATGATATCCCAATGAAAGACAGGGGAAAGCCCTGCTCTTAAGATA 1639
QY 1561 TCCATCTCTCTCTGAAACAACAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620
DB 1640 TCCATCTCTCTCTGAAACAACAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1699
QY 1621 ATATAGTGTGGTCCAGAAACCGGCATAGCCCGTTTATTTGGTTCTTACAACTAAGAG 1680
DB 1700 ATATAGTGTGGTCCAGAAACCGGCATAGCCCGTTTATTTGGTTCTTACAACTAAGAG 1759
QY 1681 AAATCTCAAGAAACAACAACCAAGATGAAATTTTGGAGCAATG---GTTTTTGGCTGC 1737
DB 1760 AAATCTCAAGAAACAACAACCAAGATGAAATTTTGGAGCAATG---GTTTTTGGCTGC 1819
QY 1738 AGGCAATAGGATAGGATTTATCTATTCAAGAAAGACTCAGACATTTCTTAAGCATGG 1797
DB 1820 AGGCAATAGGATAGGATTTATCTATTCAAGAAAGACTCAGACATTTCTTAAGCATGG 1879
QY 1798 ATCTTAATCACTTAAAGGTTTCTTCTCAAGAGATGCTCTGTGTGGGAGAGAAAGCC 1857
DB 1880 ATCTTAATCACTTAAAGGTTTCTTCTCAAGAGATGCTCTGTGTGGGAGAGAAAGCC 1939
QY 1858 CCAAGCAAGATATTAAGACAAATCCAGCTTCAATGAGCAAGAGTGTGGAGAAATCTTC 1917
DB 1940 CCAAGCAAGATATTAAGACAAATCCAGCTTCAATGAGCAAGAGTGTGGAGAAATCTTC 1999
QY 1918 CTCACAGAGAACGGCCATATTTATGTGTGTGAGATGCAAGAAATATGSCCAAGATGA 1977
DB 2000 CTCACAGAGAACGGCCATATTTATGTGTGTGAGATGCAAGAAATATGSCCAAGATGA 2059
QY 1978 CATGATGCCCTTGTGCAATTAATTAAGCAAAAGAGTGTGAGTTGAAAACTTAAGCAATG 2037
DB 2060 CATGATGCCCTTGTGCAATTAATTAAGCAAAAGAGTGTGAGTTGAAAACTTAAGCAATG 2119
```

```
QY 2038 AAAACCTGGCCACTTTAAAGAGAGAAAAAGCCCTACTTACAGATATTTGTCTATTA 2094
DB 2120 AAAACCTGGCCACTTTAAAGAGAGAAAAAGCCCTACTTACAGATATTTGTCTATTA 2176

RESULT 5
ID ADM43208
ID ADM43208 standard; cDNA; 2094 BP.
XX
AC ADM43208;
XX
DE 03-JUN-2004 (first entry)
XX
DE Human wild-type methionine synthase reductase CDS.
XX
KW Human; ss; Methionine synthase reductase polypeptide; hsmTRR; cancer;
KW cardiovascular disease; neural tube defect; hyperhomocysteinemia;
XX chromosome 5p15.2-p15.3; SNP; single nucleotide polymorphism.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..2094
XX FT /tag= a
XX FT /product= "hsmTRR"
XX FT /partial
XX FT /note= "No stop codon shown"
XX FT replace(66,A)
XX FT /tag= b
XX FT /standard_name= "single_nucleotide_polymorphism"
XX FT replace(110,A)
XX FT /tag= c
XX FT /standard_name= "single_nucleotide_polymorphism"
XX
XX US2003082676-A1.
XX
XX 01-MAY-2003.
XX
XX 10-AUG-1999; 99US-00371347.
XX
XX 16-JAN-1998; 98US-0071622P.
XX
XX 15-JAN-1999; 99US-00232028.
XX
XX (GRAY/) GRAVEL R. A.
XX (ROZE/) ROZEN R.
XX (LECL/) LECLERC D.
XX (WILS/) WILSON A.
XX (ROSE/) ROSENBLATT D.
XX
XX Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;
XX WPI; 2003-576610/54.
XX P-PSDB; ADM43207.
XX
XX New substantially pure nucleic acid encoding a mammalian methionine
XX synthase reductase polypeptide, useful for diagnosing, preventing or
XX treating conditions associated with altered methionine synthase activity,
XX e.g. cancer.
XX
XX Claim 3; SEQ ID NO 1; 26pp; English.
XX
XX The invention relates to a substantially pure nucleic acid that encodes a
XX mammalian methionine synthase reductase polypeptide, hsmTRR, or that
XX hybridizes at high stringency to a nucleic acid appearing as ADM43208 or
XX ADM43209. Also included are a non-human animal where one or both genetic
XX alleles encoding the methionine synthase reductase polypeptide are
XX mutated, an antibody that specifically binds the above methionine
XX synthase reductase polypeptide, a method of detecting the presence of the
XX above polypeptide, a method for detecting sequence variants for
XX methionine synthase reductase in a mammal, methods of creating or
XX preventing cancer (or cardiovascular disease or neural tube defects) in a
XX subject, methods of screening for a compound that modulates methionine
XX synthase reductase biological activity and a method for detecting an
```



CC increased risk of developing a neural tube defect in a mammalian embryo  
CC or foetus. The nucleic acid is useful in diagnosing, preventing or  
CC treating conditions associated with altered methionine synthase activity,  
CC such as cancer, cardiovascular disease or neural tube defects, or in  
CC screening for a compound that modulates methionine synthase reductase  
CC biological activity. Naturally occurring variants of the polypeptide are  
CC also associated with hyperhomocysteinemia. The gene for hMTRR is  
CC located on chromosome 5p15.2-p15.3. The present sequence is the coding  
CC sequence of the wild-type human hMTRR cDNA.

XX Sequence 2094 BP; 591 A; 489 C; 481 G; 533 T; 0 U; 0 Other;

Query Match 88.4%; Score 1851; DB 11; Length 2094;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 2091; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 ATGAGAGAGTTCTGTACTATATAGCTACACAGAGGAGCAAGCCATCGAGAA 60  
DB 1 ATGAGAGAGTTCTGTACTATATAGCTACACAGAGGAGCAAGCCATCGAGAA 60  
QY 61 GAAATGTGAGCAAGCTGTGTACATGATGATTTTCTGCAGATCTTCACTGTATTAAGTAA 120  
DB 61 GAAATGTGAGCAAGCTGTGTACATGATGATTTTCTGCAGATCTTCACTGTATTAAGTAA 120  
QY 121 TCCGATTAAGTATACCTTAACAAACCGAAACAGCTCTCTGTGTGTGTGTCTTACACAG 180  
DB 121 TCCGATTAAGTATACCTTAACAAACCGAAACAGCTCTCTGTGTGTGTGTCTTACACAG 180  
QY 181 GGCACCGAGACCAACCGACACAGCCCGCAAGTTGTAAAGAAATACAGAACCAACA 240  
DB 181 GGCACCGAGACCAACCGACACAGCCCGCAAGTTGTAAAGAAATACAGAACCAACA 240  
QY 241 CTGCGGTTGATTTCTTGTCTCACTGCGGTATGGGTTACTGGGCTCTCGGTATTCAGAA 300  
DB 241 CTGCGGTTGATTTCTTGTCTCACTGCGGTATGGGTTACTGGGCTCTCGGTATTCAGAA 300  
QY 301 TAAACCTACTTTTGCAATGGGGGGAATATTAATGAATTAAGACTTCAAGAGCTTGAAGCC 360  
DB 301 TAAACCTACTTTTGCAATGGGGGGAATATTAATGAATTAAGACTTCAAGAGCTTGAAGCC 360  
QY 361 CGGCATTTCTATGACACTGACATGACATGACATGACATGACATGACATGACATGACATG 420  
DB 361 CGGCATTTCTATGACACTGACATGACATGACATGACATGACATGACATGACATGACATG 420  
QY 421 CCGTGATGCTGACACTGCGGCAAGCCCTCAGAAAGATTTTAAAGTCAAGAGACAA 480  
DB 421 CCGTGATGCTGACACTGCGGCAAGCCCTCAGAAAGATTTTAAAGTCAAGAGACAA 480  
QY 481 GAGAGATTAAGTGGCGGACCTCCGCGTGGCATCACTGCAATCTTGAAGACAGACCTTGTG 540  
DB 481 GAGAGATTAAGTGGCGGACCTCCGCGTGGCATCACTGCAATCTTGAAGACAGACCTTGTG 540  
QY 541 AAGTCAGAGCTGTACACATTTGAATCTCAAGTCAAGCTTCTGAGATTCATGATTCAGAA 600  
DB 541 AAGTCAGAGCTGTACACATTTGAATCTCAAGTCAAGCTTCTGAGATTCATGATTCAGAA 600  
QY 601 AGAAGAGATTTCTAGAGTTTGAAGCAAAATGACAGTGAACCAACCAATCCAATGTTGTA 660  
DB 601 AGAAGAGATTTCTAGAGTTTGAAGCAAAATGACAGTGAACCAACCAATCCAATGTTGTA 660  
QY 661 ATTGAAGACTTTAGTGTCTCACTTACCCGTTCCGTACCCCACTCTCAAGAGCTCTCTG 720  
DB 661 ATTGAAGACTTTAGTGTCTCACTTACCCGTTCCGTACCCCACTCTCAAGAGCTCTCTG 720  
QY 721 AATATTCCTGTTTACCCCGAATATTTACAGGTATCATCTGACAGAGTCTCTTGGCCAG 780  
DB 721 AATATTCCTGTTTACCCCGAATATTTACAGGTATCATCTGACAGAGTCTCTTGGCCAG 780  
QY 781 GAGGAAAGCCAAAGTATCTGTACTTCAAGAGATCCAGTTTTCAGAGTCAATTTCAAG 840  
DB 781 GAGGAAAGCCAAAGTATCTGTACTTCAAGAGATCCAGTTTTCAGAGTCAATTTCAAG 840  
QY 841 GCACTTCACTTACTAGAAATGATGCCATTAACCACTCTGCTGATGATTTGACATT 900

DB 841 GCACTTCACTTACTAGAAATGATGCCATTAACCACTCTGCTGATGATTTGACATT 900  
QY 901 TCAATATACAGACTTTTCTATACAGCTTGAGAGATGCTTCAAGGTATCTGCCCTAACAT 960  
DB 901 TCAATATACAGACTTTTCTATACAGCTTGAGAGATGCTTCAAGGTATCTGCCCTAACAT 960  
QY 961 GATTTGAGTACAAAGCTTCTCAAGATCTGCAAGCTTGAAGATTAAGAGACCTGC 1020  
DB 961 GATTTGAGTACAAAGCTTCTCAAGATCTGCAAGCTTGAAGATTAAGAGACCTGC 1020  
QY 1021 GTCTTTTAAATTAAGGACGACCAAGAAAGAAAGAGTACCTTACCCAGCATATA 1080  
DB 1021 GTCTTTTAAATTAAGGACGACGACCAAGAAAGAAAGAGTACCTTACCCAGCATATA 1080  
QY 1081 CCGCGGAGATGTTCTCTCCAGTTCAATTTTACCTGAGTCTTGAATCCGAGCAATTCCT 1140  
DB 1081 CCGCGGAGATGTTCTCTCCAGTTCAATTTTACCTGAGTCTTGAATCCGAGCAATTCCT 1140  
QY 1141 AAAAAGCAATTTTGGAGACCTTGTGACTATACAGTGAAGTCTGAAAAGCCAG 1200  
DB 1141 AAAAAGCAATTTTGGAGACCTTGTGACTATACAGTGAAGTCTGAAAAGCCAG 1200  
QY 1201 CTACAGAGCTGTGACATTAACAAAGGGGAGCCGATTAAGCCGTTTACAGAGATGCC 1260  
DB 1201 CTACAGAGCTGTGACATTAACAAAGGGGAGCCGATTAAGCCGTTTACAGAGATGCC 1260  
QY 1261 TGTGCTGCTGTGTGATCTCTCTGCTTCTCTCTTCTTCTGCAAGCACAACATCACTC 1320  
DB 1261 TGTGCTGCTGTGTGATCTCTCTGCTTCTCTCTTCTTCTGCAAGCACAACATCACTC 1320  
QY 1321 CTGCTGAAATCTTCTTAACTTCAACCAAGCAATTCGTTGACAGTCAAGCTTCAAGTTA 1380  
DB 1321 CTGCTGAAATCTTCTTAACTTCAACCAAGCAATTCGTTGACAGTCAAGCTTCAAGTTA 1380  
QY 1381 TTTTCAACCGAAGCTCCATTTTGTCTCAACATTTGGAATTTCTGTCTACTGACACA 1440  
DB 1381 TTTTCAACCGAAGCTCCATTTTGTCTCAACATTTGGAATTTCTGTCTACTGACACA 1440  
QY 1441 ACAGAGTTCTGCGAAGGAGATATGATACAGGCTGAGCTTGTGTGTTGCTTCAAGTT 1500  
DB 1441 ACAGAGTTCTGCGAAGGAGATATGATACAGGCTGAGCTTGTGTGTTGCTTCAAGTT 1500  
QY 1501 CTTGACCAAACTATATGATCTCCATGAAGACAGGGGAAAGCCCTGCTCTTAAGATA 1560  
DB 1501 CTTGACCAAACTATATGATCTCCATGAAGACAGGGGAAAGCCCTGCTCTTAAGATA 1560  
QY 1561 TCCATCTCTCTGAAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
DB 1561 TCCATCTCTCTGAAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
QY 1621 ATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1680  
DB 1621 ATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1680  
QY 1681 AAACTCAAGAACCAACCAAGATGAAAATTTTGGAGCAATGTG---GTTTTTGGCTGC 1737  
DB 1681 AAACTCAAGAACCAACCAAGATGAAAATTTTGGAGCAATGTG---GTTTTTGGCTGC 1737  
QY 1738 AGGCATTAAGATTAAGATTAATCTATCAAGAAAGACTCAGACATTTCTTAAGCATGGG 1797  
DB 1738 AGGCATTAAGATTAAGATTAATCTATCAAGAAAGACTCAGACATTTCTTAAGCATGGG 1797  
QY 1798 ATCTTAATCACTTAAGGTTTCTCTCAAGAGATGCTCTGTTGGGAGAGAGAGCC 1857  
DB 1798 ATCTTAATCACTTAAGGTTTCTCTCAAGAGATGCTCTGTTGGGAGAGAGAGCC 1857  
QY 1858 CCAAGAAATGATACAAACCAATCAAGCTTCAAGCCAGAGAGTGGGAGAAATCTC 1917  
DB 1858 CCAAGAAATGATACAAACCAATCAAGCTTCAAGCCAGAGAGTGGGAGAAATCTC 1917  
QY 1918 CTCACAAAGATATGATCAAGCAATCAAGCTTCAAGCCAGAGAGTGGGAGAAATCTC 1977  
DB 1918 CTCACAAAGATATGATCAAGCAATCAAGCTTCAAGCCAGAGAGTGGGAGAAATCTC 1977

DB 1921 CTCGACGAGACGCGCATATTTATGTGTGATGACGAAGATATGCGCAAGATGTA 1980  
QY 1978 CATTGATCCCTTGTGCAATTAATTAAGCAAGAGGTTGAGTTGAAAACTAGAGCAATG 2037  
DB 1981 CATTGATCCCTTGTGCAATTAATTAAGCAAGAGGTTGAGTTGAAAACTAGAGCAATG 2040  
QY 2038 AAAACCTGCGCCACTTTAAAGAGAGAAAAAGCTTACCTTCAGGATATTTGGTCA 2091  
DB 2041 AAAACCTGCGCCACTTTAAAGAGAGAAAAAGCTTACCTTCAGGATATTTGGTCA 2094  
RESULT 6  
ADM43212  
ID ADM43212 standard; cDNA; 2094 BP.  
XX  
AC ADM43212;  
XX  
DT 03-JUN-2004 (first entry)  
DE Human methionine synthase reductase CDS G110A variant.  
DX  
XX  
KW Human; S9; Methionine synthase reductase polypeptide; hSMTRR; cancer;  
XX cardiovascular disease; neural tube defect; hyperhomocysteinemia;  
XX chromosome 5p15.2-p15.3; SNP; single nucleotide polymorphism.  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..2094  
FT /tag= a  
FT /product= "hSMTRR"  
FT /partial  
FT /note= "No stop codon shown"  
FT variation /replace (66,A)  
FT /standard\_name= "Single\_nucleotide\_polymorphism"  
FT variation /replace (110,G)  
FT /tag= c  
FT /standard\_name= "Single\_nucleotide\_polymorphism"  
XX  
XX US2003082676-A1.  
XX  
XX 01-MAY-2003.  
XX  
XX 10-AUG-1999; 99US-00371347.  
XX  
XX 16-JAN-1998; 98US-0071622P.  
XX 15-JAN-1999; 99US-00232028.  
XX  
XX (GRAY/) GRAVEL R A.  
XX (ROZE/) ROZEN R.  
XX (LECL/) LECLEERC D.  
XX (WILS/) WILSON A.  
XX (ROSE/) ROSENBLATT D.  
XX  
XX Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;  
XX  
XX WPI; 2003-576610/54.  
XX P-PSDB; ADM43213.  
XX  
XX  
XX New substantially pure nucleic acid encoding a mammalian methionine  
XX synthase reductase polypeptide, useful for diagnosing, preventing or  
XX treating conditions associated with altered methionine synthase activity,  
XX e.g. cancer.  
XX  
XX  
XX Disclosure; SEQ ID NO 43; 26pp; English.  
XX  
XX The invention relates to a substantially pure nucleic acid that encodes a  
XX mammalian methionine synthase reductase polypeptide, hSMTRR, or that  
XX hybridizes at high stringency to a nucleic acid appearing as ADM43208 or  
XX ADM43209. Also included are a non-human animal where one or both genetic  
XX alleles encoding the methionine synthase reductase polypeptide are  
XX mutated, an antibody that specifically binds the above methionine

CC synthase reductase polypeptide, a method of detecting the presence of the  
CC above polypeptide, a method for detecting sequence variants for  
CC methionine synthase reductase in a mammal, methods of treating or  
CC preventing cancer (or cardiovascular disease or neural tube defects) in a  
CC subject, methods of screening for a compound that modulates methionine  
CC synthase reductase biological activity and a method for detecting an  
CC increased risk of developing a neural tube defect in a mammalian embryo  
CC or foetus. The nucleic acid is useful in diagnosing, preventing or  
CC treating conditions associated with altered methionine synthase activity,  
CC such as cancer, cardiovascular disease or neural tube defects, or in  
CC screening for a compound that modulates methionine synthase reductase  
CC biological activity. Naturally occurring variants of the polypeptide are  
CC also associated with hyperhomocysteinemia. The gene for hSMTRR is  
CC located on chromosome 5p15.2-p15.3. The present sequence is the coding  
CC sequence of a variant human hSMTRR cDNA.  
XX  
XX Sequence 2094 BP; 592 A; 489 C; 480 G; 533 T; 0 U; 0 Other;  
SQ  
Query Match 86.0%; Score 1800; DB 11; Length 2094;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2090; Conservative 0; Mismatches 1; Indels 3; Gaps 1;  
QY 1 ATGAGAGGTTTCTGTACTATATGCTACACAGGAGGAGCAAGGCAATCGCAGAA 60  
DB 1 ATGAGAGGTTTCTGTACTATATGCTACACAGGAGGAGCAAGGCAATCGCAGAA 60  
QY 61 GAAATGTGTGACCAAGCTGTGTACATGATTTTCTGCAGATCTTCAACAG 120  
DB 61 GAAATGTGTGACCAAGCTGTGTACATGATTTTCTGCAGATCTTCAACAG 120  
QY 121 TCGATTAATGATGACCTTAAACCCGAAAGAGCTCTTGTGTGTGTTTCTACACG 180  
DB 121 TCGATTAATGATGACCTTAAACCCGAAAGAGCTCTTGTGTGTGTTTCTACACG 180  
QY 181 GGCACCGAGAGACCCAGCCAGACAGCCGCGAAGTTTGTAAAGAAATACAGAACCAACA 240  
DB 181 GGCACCGAGAGACCCAGCCAGACAGCCGCGAAGTTTGTAAAGAAATACAGAACCAACA 240  
QY 241 CTGCGGTTGATTTCTTGTCTCACTGCGGTATGGTTACTGGTCTCGGTATTCAGAA 300  
DB 241 CTGCGGTTGATTTCTTGTCTCACTGCGGTATGGTTACTGGTCTCGGTATTCAGAA 300  
QY 301 TACACCTACTTTTGCATGCGGAGAAATATGATTAACAGCTTCAAGAGCTTGGAGCC 360  
DB 301 TACACCTACTTTTGCATGCGGAGAAATATGATTAACAGCTTCAAGAGCTTGGAGCC 360  
QY 361 CGGCAATTTCTATGACACTGACATGACATGACATGACATGACATGACATGACATGAC 420  
DB 361 CGGCAATTTCTATGACACTGACATGACATGACATGACATGACATGACATGACATGAC 420  
QY 421 CCGTGAATGCTGACACTGCGCAGCCCTCAGAAAGCATTTTAAAGTCAAGAGAGACAA 480  
DB 421 CCGTGAATGCTGACACTGCGCAGCCCTCAGAAAGCATTTTAAAGTCAAGAGAGACAA 480  
QY 481 GAGAGATTAAGTGGGCACTCCCGGTGATCATCTGATCTTGAAGACAGACTTTGTG 540  
DB 481 GAGAGATTAAGTGGGCACTCCCGGTGATCATCTGATCTTGAAGACAGACTTTGTG 540  
QY 541 AAGTCAGAGCTGTACATGATTAATCTCAAGTCAAGCTTCTAGATTCAGATTCAGAG 600  
DB 541 AAGTCAGAGCTGTACATGATTAATCTCAAGTCAAGCTTCTAGATTCAGATTCAGAG 600  
QY 601 AGAAGGATTCGAGGTTTGAAGCAAAATGCAATGACAGCAACCAATTCATATGTTGA 660  
DB 601 AGAAGGATTCGAGGTTTGAAGCAAAATGCAATGACAGCAACCAATTCATATGTTGA 660  
QY 661 ATTGAAGATTTGAGTCCCTACCTTACCGGTGATGATCCCACTCTCAGAGCTCTGTG 720  
DB 661 ATTGAAGATTTGAGTCCCTACCTTACCGGTGATGATCCCACTCTCAGAGCTCTGTG 720  
QY 721 AATATTCCTGATTTACCCCAAGAAATATTAAGATATCAAGAGATCTCTGGCCAG 780  
DB 721 AATATTCCTGATTTACCCCAAGAAATATTAAGATATCAAGAGATCTCTGGCCAG 780

QY 781 GAGAAAGCCAGATCTGTGACTTCAGCAGATCCAGTCTTTTCAAGTGCATTTCAAG 840  
DB 781 GAGAAAGCCAGATCTGTGACTTCAGCAGATCCAGTCTTTTCAAGTGCATTTCAAG 840  
QY 841 GCAGTTCAATTCTCTGCAATGATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 900  
DB 841 GCAGTTCAATTCTCTGCAATGATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 900  
QY 901 TCAATATACAGACTTTTCTCTGCAATGATGCAATGCAATGCAATGCAATGCAATGCAAT 960  
DB 901 TCAATATACAGACTTTTCTCTGCAATGATGCAATGCAATGCAATGCAATGCAATGCAAT 960  
QY 961 GATTCTGAGGTACAAAGCTTACTTCAAAAGCTGCAAGTGAAGTAAAGAGAGCACTGC 1020  
DB 961 GATTCTGAGGTACAAAGCTTACTTCAAAAGCTGCAAGTGAAGTAAAGAGAGCACTGC 1020  
QY 1021 GTGCTTTGAAATTAAGAGAGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080  
DB 1021 GTGCTTTGAAATTAAG 1080  
QY 1081 CTTGCGGAGATGTTCTCTGCAATGATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1140  
DB 1081 CTTGCGGAGATGTTCTCTGCAATGATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1140  
QY 1141 AAAAAGGCAATTTTGGAG 1200  
DB 1141 AAAAAGGCAATTTTGGAG 1200  
QY 1201 CTACAGAGAGCTGAG 1260  
DB 1201 CTACAGAGAGAGCTGAG 1260  
QY 1261 TGTGCTGCTGTTGATGCT 1320  
DB 1261 TGTGCTGCTGTTGATGCT 1320  
QY 1321 CTGCTGCAATCTTCT 1380  
DB 1321 CTGCTGCAATCTTCT 1380  
QY 1381 TTTTACCCAG 1440  
DB 1381 TTTTACCCAG 1440  
QY 1441 ACAGAGAGCTGCGAG 1500  
DB 1441 ACAGAGAGCTGCGAG 1500  
QY 1501 CTTGAGGCAAAAT 1560  
DB 1501 CTTGAGGCAAAAT 1560  
QY 1561 TTTTACCCAG 1620  
DB 1561 TTTTACCCAG 1620  
QY 1621 ATATATGAG 1680  
DB 1621 ATATATGAG 1680  
QY 1681 AAACTCCAG 1737  
DB 1681 AAACTCCAG 1737  
QY 1738 AGGATATAG 1797  
DB 1741 AGGATATAG 1800  
QY 1798 ATCTTAATCTATTAAG 1857  
DB 1801 ATCTTAATCTATTAAG 1860

QY 1858 CCAGCAAGAT 1917  
DB 1861 CCAGCAAGAT 1920  
QY 1918 CTTGAG 1977  
DB 1921 CTTGAG 1980  
QY 1978 CATGATGCTCTGTTGCAAT 2037  
DB 1981 CATGATGCTCTGTTGCAAT 2040  
QY 2038 AAAACCTGAG 2091  
DB 2041 AAAACCTGAG 2094

RESULT 7  
ADM43209  
ID ADM43209 standard; cDNA; 2094 BP.  
XX  
AC ADM43209;  
XX  
DT 03-JUN-2004 (first entry)  
XX  
DE Human methionine synthase reductase CDS G66A variant.  
XX  
KW Human; ss; Methionine synthase reductase polypeptide; HAMTR; cancer;  
KM cardiovascular disease; neural tube defect; hyperhomocysteinemia;  
XX chromosome 5p15.2-p15.3; SNP; single nucleotide polymorphism.  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..2094  
FT /tag= a  
FT /product= "HAMTR"  
FT /partial  
FT /note= "No stop codon shown"  
FT /replace(66, G)  
FT /\*tag= b  
FT /standard\_name= "Single\_nucleotide\_polymorphism"  
FT /replace(110, A)  
FT /\*tag= c  
FT /standard\_name= "Single\_nucleotide\_polymorphism"  
XX  
PN US2003082676-A1.  
XX  
PD 01-MAY-2003.  
XX  
XX 10-AUG-1999; 99US-00371347.  
XX  
PR 16-JAN-1998; 98US-0071622P.  
PR 15-JAN-1999; 99US-00232026.  
XX  
PA (GRAY/) GRAVEL R A.  
PA (ROZE/) ROZEN R.  
PA (LECL/) LECLERC D.  
PA (WILS/) WILSON A.  
PA (ROSE/) ROSENBLATT D.  
XX  
XX Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;  
XX WPI; 2003-57610/54.  
XX DR P-PSDB; ADM43211.  
XX  
XX New substantially pure nucleic acid encoding a mammalian methionine  
XX synthase reductase polypeptide, useful for diagnosing, preventing or  
XX treating conditions associated with altered methionine synthase activity,  
XX e.g. cancer.  
XX  
XX Claim 3; SEQ ID NO 41; 26bp; English.  
XX

CC The invention relates to a substantially pure nucleic acid that encodes a  
CC mammalian methionine synthase reductase polypeptide. HmTR, or that  
CC hybridizes at high stringency to a nucleic acid appearing as ADM3208 or  
CC ADM3209. Also included are a non-human animal where one or both genetic  
CC alleles encoding the methionine synthase reductase polypeptide are  
CC mutated, an antibody that specifically binds the above methionine  
CC synthase reductase polypeptide, a method of detecting the presence of the  
CC above polypeptide, a method for detecting sequence variants for  
CC methionine synthase reductase in a mammal, methods of treating or  
CC preventing cancer (or cardiovascular disease or neural tube defects) in a  
CC subject, methods of screening for a compound that modulates methionine  
CC synthase reductase biological activity and a method for detecting an  
CC increased risk of developing a neural tube defect in a mammalian embryo  
CC or foetus. The nucleic acid is useful in diagnosing, preventing or  
CC treating conditions associated with altered methionine synthase activity,  
CC such as cancer, cardiovascular disease or neural tube defects, or in  
CC screening for a compound that modulates methionine synthase reductase  
CC biological activity. Naturally occurring variants of the polypeptide are  
CC also associated with hyperhomocysteinemia. The gene for HmTR is  
CC located on chromosome 5p15.2-p15.3. The present sequence is the coding  
CC sequence of a variant human hmTR cDNA.

XX Sequence 2094 BP; 592 A; 489 C; 480 G; 533 T; 0 U; 0 Other;

Query Match 86.0%; Score 1800; DB 11; Length 2094;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2090; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 1 ATGAGAGGTTTCTGTTACTATATGCTACACAGAGGAGCAAGGCAAAAGCCATCGAGAA 60  
DB 1 ATGAGAGGTTTCTGTTACTATATGCTACACAGAGGAGCAAGGCAAAAGCCATCGAGAA 60  
QY 61 GAAATGTGAGAGAACTGTGTACATGATTTTTCGAGATCTTCACTGATTTATGAA 120  
DB 61 GAAATGTGAGAGAACTGTGTACATGATTTTTCGAGATCTTCACTGATTTATGAA 120  
QY 121 TCCGATATGATGACCTTAAACCGAAAGAGCTCTTGTGTGTGTGTTCTACAG 180  
DB 121 TCCGATATGATGACCTTAAACCGAAAGAGCTCTTGTGTGTGTGTTCTACAG 180  
QY 121 TCCGATATGATGACCTTAAACCGAAAGAGCTCTTGTGTGTGTGTTCTACAG 180  
DB 121 TCCGATATGATGACCTTAAACCGAAAGAGCTCTTGTGTGTGTGTTCTACAG 180  
QY 181 GGCACCGAGAGCCACCCGACACAGCCCGAAGTTGTTAAGAAATACAGAACCAACA 240  
DB 181 GGCACCGAGAGCCACCCGACACAGCCCGAAGTTGTTAAGAAATACAGAACCAACA 240  
QY 241 CTGCCGTTGATTTCTTGTCTCACTGCGGTATGAGTTCTCGGTATTCAGAA 300  
DB 241 CTGCCGTTGATTTCTTGTCTCACTGCGGTATGAGTTCTCGGTATTCAGAA 300  
QY 301 TACACCTACTTTGCAATGGGGGAGATTAATTGATTAACGACTTCAAGAGCTTGAGCC 360  
DB 301 TACACCTACTTTGCAATGGGGGAGATTAATTGATTAACGACTTCAAGAGCTTGAGCC 360  
QY 361 CGGCAATTTCTATGACACTGACATGCAATGATGATGATTTAGAACTTGTGTTAG 420  
DB 361 CGGCAATTTCTATGACACTGACATGCAATGATGATGATTTAGAACTTGTGTTAG 420  
QY 421 CGGTGATTTGCTGAGCTTGGCCAGCCCTCAGAAAGATTTTATGTCAGAGAGACA 480  
DB 421 CGGTGATTTGCTGAGCTTGGCCAGCCCTCAGAAAGATTTTATGTCAGAGAGACA 480  
QY 481 GAGGAGATTAAGTGGGCACTCCCGGTGGATCACTGCACTCTTGAGAGACAGACTTGG 540  
DB 481 GAGGAGATTAAGTGGGCACTCCCGGTGGATCACTGCACTCTTGAGAGACAGACTTGG 540  
QY 541 AAGTCAGAGCTGTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
DB 541 AAGTCAGAGCTGTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
QY 601 AGAAAGATTTCTGAGTTTGAAGCAAAATGCAATGCAAGCAACCAATTCATGTTGTA 660  
DB 601 AGAAAGATTTCTGAGTTTGAAGCAAAATGCAATGCAAGCAACCAATTCATGTTGTA 660  
QY 661 ATGAAGACTTGAAGTCTCACTTACCCGTTGGTACCCCACTCTCAAGGCTCTCTG 720

DB 661 ATGAAGACTTGAAGTCTCACTTACCCGTTGGTACCCCACTCTCAAGGCTCTCTG 720  
QY 721 AATATTCCTGTTTACCCCGCAATATTTACAGATACATCTGACAGAGCTTGGCCAG 780  
DB 721 AATATTCCTGTTTACCCCGCAATATTTACAGATACATCTGACAGAGCTTGGCCAG 780  
QY 781 GAGAAAGCAAGTATCTGTGACTTACAGAGATTCAGATTTTCAAGTGCATTTCAAG 840  
DB 781 GAGAAAGCAAGTATCTGTGACTTACAGAGATTCAGATTTTCAAGTGCATTTCAAG 840  
QY 841 GCAATTCAGATTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
DB 841 GCAATTCAGATTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
QY 901 TCAATACAGATTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
DB 901 TCAATACAGATTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
QY 961 GATTCGAGATTCAGAGCTTACCTCAAGAGCTGAGATTCAGATTCAGAGCTGAG 1020  
DB 961 GATTCGAGATTCAGAGCTTACCTCAAGAGCTGAGATTCAGATTCAGAGCTGAG 1020  
QY 1021 GTCTTTGAAATTAAGGACAGACCAAGAAAGAGGCTTACCTTACCCAGCATATA 1080  
DB 1021 GTCTTTGAAATTAAGGACAGACCAAGAAAGAGGCTTACCTTACCCAGCATATA 1080  
QY 1081 CCGGAGGATGTTCTCTCAAGTTCAATTTTACCTGATCTTGAATTCAGACATTTCT 1140  
DB 1081 CCGGAGGATGTTCTCTCAAGTTCAATTTTACCTGATCTTGAATTCAGACATTTCT 1140  
QY 1141 AAAAAAGCATTTTTCGAGCCCTTGTGACATTAACAGTACAGTGTGTAAGAGCCAG 1200  
DB 1141 AAAAAAGCATTTTTCGAGCCCTTGTGACATTAACAGTACAGTGTGTAAGAGCCAG 1200  
QY 1201 CTACAGAGCTGTGACATTAACAGAGGAGCCGATTAATAGCCGTTTGTACAGATGCC 1260  
DB 1201 CTACAGAGCTGTGACATTAACAGAGGAGCCGATTAATAGCCGTTTGTACAGATGCC 1260  
QY 1261 TGTGCTGTTGTGATGATCTCTCTGCTGCTTCCCTTCTGACAGCAGCACTCACTC 1320  
DB 1261 TGTGCTGTTGTGATGATCTCTCTGCTGCTTCCCTTCTGACAGCAGCACTCACTC 1320  
QY 1321 CTGCTGAACTCTTCTTAACTTGAACCAATGATGATGATGATGATGATGATGATGAT 1380  
DB 1321 CTGCTGAACTCTTCTTAACTTGAACCAATGATGATGATGATGATGATGATGATGAT 1380  
QY 1381 TTTCAACCGAGAAAGCTCAATTTGCTTCAATTTGATGATGATGATGATGATGATGAT 1440  
DB 1381 TTTCAACCGAGAAAGCTCAATTTGCTTCAATTTGATGATGATGATGATGATGATGAT 1440  
QY 1441 ACAGAGGTTCTGCGAAGGAGATATGATGATGATGATGATGATGATGATGATGATGAT 1500  
DB 1441 ACAGAGGTTCTGCGAAGGAGATATGATGATGATGATGATGATGATGATGATGATGAT 1500  
QY 1501 CTTGAGCCAAATATCATGATCCCATGAAAGCAGCGGAAAGCCCTGCTCTTAAGATA 1560  
DB 1501 CTTGAGCCAAATATCATGATCCCATGAAAGCAGCGGAAAGCCCTGCTCTTAAGATA 1560  
QY 1561 TCCATCTCTCTGAGCAACAATTTCTTCACTTACAGATGAGCCCTCAATCCCATC 1620  
DB 1561 TCCATCTCTCTGAGCAACAATTTCTTCACTTACAGATGAGCCCTCAATCCCATC 1620  
QY 1621 ATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
DB 1621 ATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
QY 1681 AACTCCAAAGCAACACCCAGATGAAATTTTGGAGCAATGAG--GTTTGGCTG 1737  
DB 1681 AACTCCAAAGCAACACCCAGATGAAATTTTGGAGCAATGAG--GTTTGGCTG 1737  
QY 1738 AGGCAATGAAGATGAGATTTATCTATTCAGAAAGAGCTCAGCATTTCTTAAAGATGAG 1797  
DB 1738 AGGCAATGAAGATGAGATTTATCTATTCAGAAAGAGCTCAGCATTTCTTAAAGATGAG 1797



QY 961 GATTCTGAGTACAAAGCCTACTCCAAAGCTGACCTGAAAGATTAAGAGACATCC 1020  
DB 1040 GATTCTGAGTACAAAGCCTACTCCAAAGCTGACCTGAAAGATTAAGAGACATCC 1099  
QY 1021 GTCTTTTAAATTAAGGACAGACAAAGAAAGAGACTACTTACCCAGCATATA 1080  
DB 1100 GTCTTTTAAATTAAGGACAGACAAAGAAAGAGACTACTTACCCAGCATATA 1159  
QY 1081 CCTGGCGGATGTTCTCTCAGATTCAATTTTACCTGAGTGTGAAATCCGAGCAATTCCT 1140  
DB 1160 CCGGGGAGATGTTCTCTCAGATTCAATTTTACCTGAGTGTGAAATCCGAGCAATTCCT 1219  
QY 1141 AAAAAGGCAATTTTGGAGACCCCTTGGAGACTATACAGAGACAGTCTGAAAAAGCCAGG 1200  
DB 1220 AAAAAGGCAATTTTGGAGACCCCTTGGAGACTATACAGAGACAGTCTGAAAAAGCCAGG 1279  
QY 1201 CTACAGAGACTGTGAGTAAACAAAGGGGAGCCGATTAATAGCCGCTTTGTAGAGATGCC 1260  
DB 1280 CTACAGAGACTGTGAGTAAACAAAGGGGAGCCGATTAATAGCCGCTTTGTAGAGATGCC 1339  
QY 1261 TGTGCTGCTTGTGAGATCTCTCTCTGCTTCTCTTCTTCCAGCCACACTCACTCTC 1320  
DB 1340 TGTGCTGCTTGTGAGATCTCTCTCTGCTTCTCTTCTTCCAGCCACACTCACTCTC 1399  
QY 1321 CTGCTGGAACATTTCTTAACTTCAACCCAGACATATTCGTGTGAGAGCTCAAGTTTA 1380  
DB 1400 CTGCTGGAACATTTCTTAACTTCAACCCAGACATATTCGTGTGAGAGCTCAAGTTTA 1459  
QY 1381 TTTCAACCGAAGAGCTCAATTTGTCTTCAACATTTGTGAATTTCTGTCTACCTGACCA 1440  
DB 1460 TTTCAACCGAAGAGCTCAATTTGTCTTCAACATTTGTGAATTTCTGTCTACCTGACCA 1519  
QY 1441 ACAAGGTTCTGCGAAGGAGATATGATACAGCTGAGCTGAGCTTGTGTGCTTCAAGTT 1500  
DB 1520 ACAAGGTTCTGCGAAGGAGATATGATACAGCTGAGCTGAGCTTGTGTGCTTCAAGTT 1579  
QY 1501 CTTCAGCCAAACATCATGATCCCATGAAAGACAGGGGAAAGCCCTGGCTCTTAAGATA 1560  
DB 1580 CTTCAGCCAAACATCATGATCCCATGAAAGACAGGGGAAAGCCCTGGCTCTTAAGATA 1639  
QY 1561 TCCATCTCTCTGGAACAACAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
DB 1640 TCCATCTCTCTGGAACAACAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1699  
QY 1621 ATATATGTTGGTTCAGAAACCGGACATAGCCCGTTTATTTGGTTCCTACAACTAGAGAG 1680  
DB 1700 ATATATGTTGGTTCAGAAACCGGACATAGCCCGTTTATTTGGTTCCTACAACTAGAGAG 1759  
QY 1681 AAATCCAGAACACACCCAGATGGAATTTTGGAGCAATGTG---GTTTTTGGCTGC 1737  
DB 1760 AAATCCAGAACACACCCAGATGGAATTTTGGAGCAATGTG---GTTTTTGGCTGC 1819  
QY 1738 AGGCATTAAGGATGGAATTAATCTATTCAGAAAGAGCTGAGCATTTCTTAAGCATGGG 1797  
DB 1820 AGGCATTAAGGATGGAATTAATCTATTCAGAAAGAGCTGAGCATTTCTTAAGCATGGG 1879  
QY 1798 ATCTTAATCTATTAAGGTTTCTCTCAAGAGATGCTCTGTTGGGAGAGAGAGGCC 1857  
DB 1880 ATCTTAATCTATTAAGGTTTCTCTCAAGAGATGCTCTGTTGGGAGAGAGAGGCC 1939  
QY 1858 CCGAGAAAGTATATACAGACAAATCCAGCTTCAATGCGAGAGAGAGAGATCTC 1917  
DB 1940 CCGAGAAAGTATATACAGACAAATCCAGCTTCAATGCGAGAGAGAGAGATCTC 1999  
QY 1918 CTCAGAGAGAGGCGCATTTTATGTTGTGAGATGCAAGATATATGCGCAAGATGTA 1977  
DB 2000 CTCAGAGAGAGGCGCATTTTATGTTGTGAGATGCAAGATATATGCGCAAGATGTA 2059  
QY 1978 CATGATGCCCTTGTGCAATATATAGCAAGAGGTTGAGATTGAAAACTAGAAAGCATG 2037  
DB 2060 CATGATGCCCTTGTGCAATATATAGCAAGAGGTTGAGATTGAAAACTAGAAAGCATG 2119

QY 2038 AAAACCTGGCCACTTTAAAGAGAAAAACCTACTCTTACAGATATTTGTCAATA 2094  
DB 2120 AAAACCTGGCCACTTTAAAGAGAAAAACCTACTCTTACAGATATTTGTCAATA 2176

RESULT 9  
ADQ87538  
ID ADQ87538 standard; cDNA; 3270 BP.

AC ADQ87538;  
XX  
XX  
DT 07-OCT-2004 (first entry)  
XX  
DE Human tumour-associated antigenic target (TAT) cDNA sequence #4416.  
XX human; tumour-associated antigenic target; TAT; cytosolic; gene therapy;  
KW cancer; cell proliferative disorder; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN MO2004060270-A2.  
XX  
PD 22-JUL-2004.  
XX  
PF 15-OCT-2003; 2003WO-US029126.  
PR 18-OCT-2002; 2002US-0418988P.  
XX  
XX (GETH) GENENTECH INC.  
PA (WUTD/) WU T D.  
PA (ZHOU/) ZHOU Y.  
XX  
PI Wu TD, Zhou Y;  
XX  
XX MPI; 2004-534300/51.  
XX  
PT New nucleic acid molecule and encoded polypeptide, for diagnosing,  
PT preventing or treating cell proliferative disorders such as cancer.  
PS  
PS Claim 1; SEQ ID NO 4416; 5504dp; English.

The present invention describes an isolated tumour-associated antigenic target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of (a); (c) the complement of (a) or (b); (d) a sequence that has 80% sequence identity to (a)-(c); or (e) a sequence that hybridizes to (a)-(c). Also described: (1) an expression vector comprising the above nucleic acid; (2) a host cell comprising the above expression vector; (3) a process for producing a polypeptide; (4) an isolated polypeptide comprising: (a) an amino acid sequence encoded by any of the above nucleotide sequences; (b) an amino acid sequence encoded by the full-length coding region of the above nucleotide sequences; or (c) a sequence having at least 80% identical to (a) or (b); (5) a chimeric polypeptide comprising the above polypeptide fused to a heterologous polypeptide; (6) an isolated antibody that binds to the above polypeptide; (7) a process for producing the antibody; (8) an isolated oligopeptide that binds to the above polypeptide; (9) a tumour-associated antigenic target (TAT) binding organic molecule that binds to the above polypeptide; (10) a composition of matter comprising the above (chimeric) polypeptide, antibody, oligopeptide or TAT binding organic molecule, in combination with a carrier; (11) an article of manufacture comprising a container and the composition of matter contained within the container; (12) methods of inhibiting the growth of a cell that expresses the above protein, where the growth of the cell is at least in part dependent upon a growth potentiating effect of the above protein; (13) a method of therapeutically treating a mammal having a cancerous tumour comprising cells that express the above protein; (14) a method of determining the presence of a protein in a sample suspected of containing the protein described above; (15) methods of diagnosing the presence of a tumour in a mammal; (16) a method for treating or preventing a cell proliferative disorder associated with increased expression or activity of the above protein; and (17) a method of binding an antibody, oligopeptide or organic molecule to a cell that expresses the protein described above.







```
DB 2092 CATGAAGCCCTTGTGCAAAATAATACCAAGGTTGAGTTGAAAACTAGAACCAATG 2151
QY 2038 AAAACCCCTGCGCACTTTAAAGAGAAAACGCTACCTCAGAGATTTTGGTCATTA 2094
DB 2152 AAAACCCCTGCGCACTTTAAAGAGAAAACGCTACCTCAGAGATTTTGGTCATTA 2208

RESULT 10
ADM43214
ID ADM43214 standard; cDNA; 2091 BP.
XX
AC ADM43214;
XX
DT 03-JUN-2004 (first entry)
XX
DE Human methionine synthase reductase CDS del 1675-1678 variant.
XX
KW Human; ss; Methionine synthase reductase polypeptide; hsmtrr; cancer;
KM cardiovascular disease; neural tube defect; hyperhomocysteinemia;
XX chromosome 5p15.2-p15.3; SNP; single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..2091
FT /tag= a
FT /product= "hsmtrrdelR59"
FT /partial
FT /note= "No stop codon shown"
FT /replace (66,A)
FT /tag= b
FT /standard name= "single_nucleotide polymorphism"
FT /replace (110,A)
FT /tag= c
FT /standard name= "single_nucleotide polymorphism"
FT /replace (1675,AGAG)
FT /tag= d
FT variation
XX
PN US2003082676-A1.
XX
PD 01-MAY-2003.
XX
PF 10-AUG-1999; 99US-00371347.
XX
PR 16-JAN-1998; 98US-0071622P.
XX 15-JAN-1999; 99US-00232028.
XX
PA (GRAV/) GRAVEL R A.
PA (ROZE/) ROZEN R.
PA (LECL/) LECLERC D.
PA (WILS/) WILSON A.
PA (ROSE/) ROSENBLATT D.
XX
PI Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;
XX WPI; 2003-576610/54.
XX F-PSDB; ADM43215.
XX
PT New substantially pure nucleic acid encoding a mammalian methionine
PT synthase reductase polypeptide, useful for diagnosing, preventing or
PT treating conditions associated with altered methionine synthase activity,
PT e.g. cancer.
XX
PS Disclosure; SEQ ID NO 47; 26pp; English.
XX
XX The invention relates to a substantially pure nucleic acid that encodes a
XX mammalian methionine synthase reductase polypeptide, hsmtrr, or that
XX hybridizes at high stringency to a nucleic acid appearing as ADM43208 or
XX ADM43209. Also included are a non-human animal where one or both genetic
XX alleles encoding the methionine synthase reductase polypeptide are
XX mutated, an antibody that specifically binds the above methionine
XX synthase reductase polypeptide, a method of detecting the presence of the
```

```
CC above polypeptide, a method for detecting sequence variants for
CC methionine synthase reductase in a mammal, methods of treating or
CC preventing cancer (or cardiovascular disease or neural tube defects) in a
CC subject, methods of screening for a compound that modulates methionine
CC synthase reductase biological activity and a method for detecting an
CC increased risk of developing a neural tube defect in a mammalian embryo
CC or foetus. The nucleic acid is useful in diagnosing, preventing or
CC treating conditions associated with altered methionine synthase activity,
CC such as cancer, cardiovascular disease or neural tube defects, or in
CC screening for a compound that modulates methionine synthase reductase
CC biological activity. Naturally occurring variants of the polypeptide are
CC also associated with hyperhomocysteinemia. The gene for hsmtrr is
CC located on chromosome 5p15.2-p15.3. The present sequence is the coding
CC sequence of a variant human hsmtrr cDNA.
XX
SQ Sequence 2091 BP; 589 A; 489 C; 480 G; 533 T; 0 U; 0 Other;
Query Match 79.9%; Score 1674; DB 11; Length 2091;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGAGAGGTTCTGTACTATATGCTACAGCAGGAGCAGGCAAGGCCATCGAGAA 60
DB 1 ATGAGAGAGGTTCTGTACTATATGCTACAGCAGGAGCAGGCAAGGCCATCGAGAA 60
QY 61 GAAATGTGAGCAAGCTGTGTACATGATTTTCTGAGATCTTCACTGATTAAGTGA 120
DB 61 GAAATGTGAGCAAGCTGTGTACATGATTTTCTGAGATCTTCACTGATTAAGTGA 120
QY 121 TCCGATTAAGTATGACTTAAACCGAAGCAGCTCTTGTGTGTGTGTTCTACACG 180
DB 121 TCCGATTAAGTATGACTTAAACCGAAGCAGCTCTTGTGTGTGTGTTCTACACG 180
QY 121 TCCGATTAAGTATGACTTAAACCGAAGCAGCTCTTGTGTGTGTGTTCTACACG 180
DB 121 TCCGATTAAGTATGACTTAAACCGAAGCAGCTCTTGTGTGTGTGTTCTACACG 180
QY 181 GGCACCGGAGACCCACCCGACACAGCCGCAAGTTGTGTAAGAAATACGAACCAACA 240
DB 181 GGCACCGGAGACCCACCCGACACAGCCGCAAGTTGTGTAAGAAATACGAACCAACA 240
QY 181 GGCACCGGAGACCCACCCGACACAGCCGCAAGTTGTGTAAGAAATACGAACCAACA 240
DB 181 GGCACCGGAGACCCACCCGACACAGCCGCAAGTTGTGTAAGAAATACGAACCAACA 240
QY 241 CTGCCGTTGATTTCTTCTGCTACCTGCGGTATGAGTTTCTGGGTCTCGGTATTCAGAA 300
DB 241 CTGCCGTTGATTTCTTCTGCTACCTGCGGTATGAGTTTCTGGGTCTCGGTATTCAGAA 300
QY 301 TACACCTACTTTTGCAATGGGGGGAAGATATGATTAACGACTTCAGAGCTTGGAGCC 360
DB 301 TACACCTACTTTTGCAATGGGGGGAAGATATGATTAACGACTTCAGAGCTTGGAGCC 360
QY 301 TACACCTACTTTTGCAATGGGGGGAAGATATGATTAACGACTTCAGAGCTTGGAGCC 360
DB 301 TACACCTACTTTTGCAATGGGGGGAAGATATGATTAACGACTTCAGAGCTTGGAGCC 360
QY 361 CGGCATTTCTATGACACTGGAATGACATGACTGTGTAGTTTGAACCTTGTGTTGAG 420
DB 361 CGGCATTTCTATGACACTGGAATGACATGACTGTGTAGTTTGAACCTTGTGTTGAG 420
QY 421 CCGTGGATTGCTGGACTGCGCAGCCCTCAGAAAGCATTTTAGGTCAGAGAGACAA 480
DB 421 CCGTGGATTGCTGGACTGCGCAGCCCTCAGAAAGCATTTTAGGTCAGAGAGACAA 480
QY 481 GAGAGATTAAGTGGGCACTCCCGGTGATCATCTGCATCTTGAAGACAGACTTGTG 540
DB 481 GAGAGATTAAGTGGGCACTCCCGGTGATCATCTGCATCTTGAAGACAGACTTGTG 540
QY 541 AAGTCAGAGCTGCTACATTTGAATCTCAAGTCGAGCTTCTGAGATTCATGATTCAGGA 600
DB 541 AAGTCAGAGCTGCTACATTTGAATCTCAAGTCGAGCTTCTGAGATTCATGATTCAGGA 600
QY 601 AAGTCAAGCTGCTACATTTGAATCTCAAGTCGAGCTTCTGAGATTCATGATTCAGGA 660
DB 601 AAGTCAAGCTGCTACATTTGAATCTCAAGTCGAGCTTCTGAGATTCATGATTCAGGA 660
QY 601 AAGTCAAGCTGCTACATTTGAATCTCAAGTCGAGCTTCTGAGATTCATGATTCAGGA 660
DB 601 AAGTCAAGCTGCTACATTTGAATCTCAAGTCGAGCTTCTGAGATTCATGATTCAGGA 660
QY 661 ATTGAAGACTTGAAGCTCCTACCTACCCGTTGATACCCCACTCACAAGCTTCTG 720
DB 661 ATTGAAGACTTGAAGCTCCTACCTACCCGTTGATACCCCACTCACAAGCTTCTG 720
QY 721 AATATTCTGTTTACCCCAAGATATTTTACAGTATCATCTGACAGAGTCTTGGCCAG 780
DB 721 AATATTCTGTTTACCCCAAGATATTTTACAGTATCATCTGACAGAGTCTTGGCCAG 780
```



```
QY 421 CCGTGAATTCGTCGACTCTGCGCAGGCCCTCAGAAAACATTTTAAAGTCAGACAGAGACAA 480
DB 500 CCGTGAATTCGTCGACTCTGCGCAGGCCCTCAGAAAACATTTTAAAGTCAGACAGAGACAA 559
QY 481 GAGGAGATTAAGTGGGCGACTCCCGGTGGATCACTTGACCTTTGAGGACAGACCTTGG 540
DB 560 GAGGAGATTAAGTGGGCGACTCCCGGTGGATCACTTGACCTTTGAGGACAGACCTTGG 619
QY 541 AAGTCAGAGCTGCTACACATTTGAATCTCAAGTCAGCTTGAAGATTCAGATTCAGGA 600
DB 620 AAGTCAGAGCTGCTACACATTTGAATCTCAAGTCAGCTTGAAGATTCAGATTCAGGA 679
QY 601 AGAAAAGATTCGAGGTTTGAAGCAAAAATGCAAGCAAGCAACCAATCAATGTTGTA 660
DB 680 AGAAAAGATTCGAGGTTTGAAGCAAAAATGCAAGCAAGCAACCAATCAATGTTGTA 739
QY 661 ATTGAAGACTTGAAGTCTGACTTACCCGTTGGTATCCCACTCTCAAGCTCTCTG 720
DB 740 ATTGAAGACTTGAAGTCTGACTTACCCGTTGGTATCCCACTCTCAAGCTCTCTG 799
QY 721 AATATTCCTGCTTACCCCGAATATTTACAGGTACATCTGACAGAGTCTTGCGCAG 780
DB 800 AATATTCCTGCTTACCCCGAATATTTACAGGTACATCTGACAGAGTCTTGCGCAG 859
QY 781 GAGGAAAGCCAAAGTATCTGTGACTTCAGAGATCCAGTTTTCAGTGCCTTCAAG 840
DB 860 GAGGAAAGCCAAAGTATCTGTGACTTCAGAGATCCAGTTTTCAGTGCCTTCAAG 919
QY 841 GCAATTCAACTTACATGCAATGATGCAATTAACCACTCTGCTGATGAATTTGACAT 900
DB 920 GCAATTCAACTTACATGCAATGATGCAATTAACCACTCTGCTGATGAATTTGACAT 979
QY 901 TGAATATGACACTTTTCTATCAGAGTGAAGTCTTCAAGGATTCAGCTTACAGT 960
DB 980 TGAATATGACACTTTTCTATCAGAGTGAAGTCTTCAAGGATTCAGCTTACAGT 1039
QY 961 GATTCGAGGTACAAAGCTTACCTCAAAAGCTGACCTTGAAGTAAAGAGACACTGC 1020
DB 1040 GATTCGAGGTACAAAGCTTACCTCAAAAGCTGACCTTGAAGTAAAGAGACACTGC 1099
QY 1021 GTCTTTTGAATAAAGGAGACACAAAGAAAGAGCTACCTTACCCGACATATA 1080
DB 1100 GTCTTTTGAATAAAGGAGACACAAAGAAAGAGCTACCTTACCCGACATATA 1159
QY 1081 CCGTGGGAGATGTTCTCTCAGATTCATTTTACCTGCTGCTTGAATCCGAGAAATTCCT 1140
DB 1160 CCGTGGGAGATGTTCTCTCAGATTCATTTTACCTGCTGCTTGAATCCGAGAAATTCCT 1219
QY 1141 AAAAAGGACATTTTGGGAGCCCTTGTGACATAACAGTGAAGTGTGAAAAAGGCGAG 1200
DB 1220 AAAAAGGACATTTTGGGAGCCCTTGTGACATAACAGTGAAGTGTGAAAAAGGCGAG 1279
QY 1201 CTACAGAGCTGTGACATTAACAGAGGGGAGCGATTAATAGCGCTTTGTACAGATGCG 1260
DB 1280 CTACAGAGCTGTGACATTAACAGAGGGGAGCGATTAATAGCGCTTTGTACAGATGCG 1339
QY 1261 TGTGCTGCTTGTGATCTCTCTGCTTTCCTTTTCCCTTTCGACGACCACTCACTCTC 1320
DB 1340 TGTGCTGCTTGTGATCTCTCTGCTTTCCTTTTCCCTTTCGACGACCACTCACTCTC 1399
QY 1321 CTGCTGGAACATCTTCTAAACTTCAACCAAGCAATTCGATGTCAGAGCTCAAGTTTA 1380
DB 1400 CTGCTGGAACATCTTCTAAACTTCAACCAAGCAATTCGATGTCAGAGCTCAAGTTTA 1459
QY 1381 TTTGACCCAGAAAGCTCAATTTTGTCTTCAACATTTGAGATTTTGTCTACCTGACACA 1440
DB 1460 TTTGACCCAGAAAGCTCAATTTTGTCTTCAACATTTGAGATTTTGTCTACCTGACACA 1519
QY 1441 ACAGAGTTCGCGGAGGAGATGTAACAGCTGCTGCTGCTTGTGCTTCAAGTT 1500
DB 1520 ACAGAGTTCGCGGAGGAGATGTAACAGCTGCTGCTGCTTGTGCTTCAAGTT 1579
QY 1501 CTTCAGCCAAACATACATGATCCCATGAAGACAGGGGAAAGCCTGCGCTCTAAGATA 1560
```

```
DB 1580 CTTACGCCAAACATACATGATCCCATGAAGACAGCGGAAAGCCCTGCTCTAAGATA 1639
QY 1561 TCCATCTCTCCCTGGAACAAATTTCTTCCACTTACAGAGATGACCCCTCAATCCCATC 1620
DB 1640 TCCATCTCTCTCCCTGGAACAAATTTCTTCCACTTACAGAGATGACCCCTCAATCCCATC 1699
QY 1621 ATAAATGTTGGTCCAGAAACCGGCAT 1646
DB 1700 ATAAATGTTGGTCCAGAAACCGGCAT 1725

RESULT 12
AAAS8976
ID AAAS8976 standard; DNA; 3255 BP.
XX
AC AAAS8976;
XX
DT 07-NOV-2000 (first entry)
XX
DE A human methionine synthase reductase DNA sequence with polymorphism.
XX
KW Human; methionine synthase reductase; MTRR; cancer;
KW cardiovascular disease; Down's Syndrome; neural tube defect;
KW premature coronary artery disease; ss.
XX
OS Homo sapiens.
XX
PN WO200042196-A2.
XX
PD 20-JUL-2000.
XX
PF 14-JAN-2000; 2000MO-1B000209.
XX
PR 15-JAN-1999; 99US-00232028.
XX
PR 10-AUG-1999; 99US-00371347.
XX
PA (UWMC-) UNIV MCGILL.
XX
PI Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;
XX
DR WPI; 2000-466131/40.
XX
PT Mammalian methionine synthase reductase nucleic acid used for detecting
PT an increased risk of developing a neural tube defect, Down's Syndrome or
PT cardiovascular disease in a mammalian embryo or fetus.
XX
PS Claim 7; Page; 85pp; English.
XX
CC The present sequence represents a human methionine synthase reductase
CC (MTRR) DNA sequence, with a polymorphism comprising of a deletion of
CC nucleotides 1675-1678. Inhibitors of MTRR polypeptide and polynucleotide
CC are used for treating or preventing cancer, cardiovascular disease,
CC Down's Syndrome or neural tube defects in a subject. The cardiovascular
CC disease is premature coronary artery disease. The compounds are detected
CC by methods which screen for modulators of MTRR biological activity. MTRR
CC polypeptide or nucleic acid is examined for the presence of a
CC polymorphism in the parents or the embryo or foetus, and the information
CC used for detecting an increased risk of an embryo or foetus developing
CC cancer, cardiovascular disease, Down's Syndrome or neural tube defects.
CC note: the present sequence does not appear in the specification; it was
CC created using information provided.
XX
SQ Sequence 3255 BP; 942 A; 704 C; 663 G; 946 T; 0 U; 0 Other;

Query Match 73.7%; Score 1544; DB 3; Length 3255;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1594; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGAGAGTTTCTGTACTATATGCTACACAGGAGGACAGGCAAGGCCATTCGAGAA 60
DB 80 ATGAGAGAGTTTCTGTACTATATGCTACACAGGAGGAGGCAAGGCCATTCGAGAA 139
```

QY 61 GAAATGTGTAGCAAGCTGTGTACATGATTTTCTGCAGATCTTCACTGATTTAGTGAA 120  
Db 140 GAAATGTGTAGCAAGCTGTGTACATGATTTTCTGCAGATCTTCACTGATTTAGTGAA 199  
QY 121 TCCGATTAAGTATGACCTTAACAAACCGAAACAGCTCCTCTGTGTGTGTGTTTCTACACG 180  
Db 200 TCCGATTAAGTATGACCTTAACAAACCGAAACAGCTCCTCTGTGTGTGTGTTTCTACACG 259  
QY 181 GGCACCGGAGACCCACCGACACAGCCCGCAAGTTTGTAAAGAAATACAAACCA 240  
Db 260 GGCACCGGAGACCCACCGACACAGCCCGCAAGTTTGTAAAGAAATACAAACCA 319  
QY 241 CTGCGGTTTATTTCTTGTGTCACTGTGCGGTATGAGGTACTGTGCTCGGTGTTCAGAA 300  
Db 320 CTGCGGTTTATTTCTTGTGTCACTGTGCGGTATGAGGTACTGTGCTCGGTGTTCAGAA 379  
QY 301 TACACCTACTTTTGCATATGGGGGGAATATTAATGAATTAAGCACTTCAAGAGCTTGGAGC 360  
Db 380 TACACCTACTTTTGCATATGGGGGGAATATTAATGAATTAAGCACTTCAAGAGCTTGGAGC 439  
QY 361 CGGCAATTTCTATGACACTGTGACATGACATGACTGTGTAGGTTTAAACTTGTGTGAG 420  
Db 440 CGGCAATTTCTATGACACTGTGACATGACATGACTGTGTAGGTTTAAACTTGTGTGAG 499  
QY 421 CCGTGTATTTGCTGACCTGTGGCCAGCCCTCAAGAAAGATTTTATAGTACAGAGACAA 480  
Db 500 CCGTGTATTTGCTGACCTGTGGCCAGCCCTCAAGAAAGATTTTATAGTACAGAGACAA 559  
QY 481 GAGGAGATTAAGTGGCGGCACTCCGCGTGGCATCACCTGATCTTTGAGAGACAGACTTGTG 540  
Db 560 GAGGAGATTAAGTGGCGGCACTCCGCGTGGCATCACCTGATCTTTGAGAGACAGACTTGTG 619  
QY 541 AAGTCAGAGCTGTACACATTTGAATCTCAAGTCGAGCTTGTGAGTTTGATTCAGAA 600  
Db 620 AAGTCAGAGCTGTACACATTTGAATCTCAAGTCGAGCTTGTGAGTTTGATTCAGAA 679  
QY 601 AGAAAGATTTCTGAGTTTGAAGCAAAATGCACTGAACAGCAACCAATCATGTTGTA 660  
Db 680 AGAAAGATTTCTGAGTTTGAAGCAAAATGCACTGAACAGCAACCAATCATGTTGTA 729  
QY 661 ATTGAAGACTTTGAGTCTCTACCTTACCCGTTCCGTAACCCCACTTCAAGCCCTCTCTG 720  
Db 740 ATTGAAGACTTTGAGTCTCTACCTTACCCGTTCCGTAACCCCACTTCAAGCCCTCTCTG 799  
QY 721 AATATTCCTGTGTTTACCCCAAGAAATTTTACAGGTATCATCTGACAGAGTCTCTTGGCAG 780  
Db 800 AATATTCCTGTGTTTACCCCAAGAAATTTTACAGGTATCATCTGACAGAGTCTCTTGGCAG 859  
QY 781 GAGGAAAGCCAAAGTATCTGTGACTTTCAGCAGATTCAGATTTTCAAGTCCCAATTTCAAAG 840  
Db 860 GAGGAAAGCCAAAGTATCTGTGACTTTCAGCAGATTCAGATTTTCAAGTCCCAATTTCAAAG 919  
QY 841 GCAGTTCAACTTACTACGAATGATGATCAATTAACCACTGTGCTGTGATTAATGACATT 900  
Db 920 GCAGTTCAACTTACTACGAATGATGATCAATTAACCACTGTGCTGTGATTAATGACATT 979  
QY 901 TCAAAATACAGACTTTTCTATTCAGCTGTGAGATCCCTTCAAGCGTATCTGCCCTTAACAGT 960  
Db 980 TCAAAATACAGACTTTTCTATTCAGCTGTGAGATCCCTTCAAGCGTATCTGCCCTTAACAGT 1039  
QY 961 GATTCTAGGTATCAAAAGCTTACTCCAAAGACTGTGAGCTGTGAAGTAAAGAGAGCACTGC 1020  
Db 1040 GATTCTAGGTATCAAAAGCTTACTCCAAAGACTGTGAGCTGTGAAGTAAAGAGAGCACTGC 1099  
QY 1021 GTCCTTTTGAATAATTAAGGACGACACAAAGAAAGAGCTTACTTACCCAGCATATA 1080  
Db 1100 GTCCTTTTGAATAATTAAGGACGACACAAAGAAAGAGCTTACTTACCCAGCATATA 1159  
QY 1081 CTGCGGGAATGTTCTCTCAAGTTCAATTTTACCTGTGTCTTGAATTCGAGCAATTCCT 1140  
Db 1160 CTGCGGGAATGTTCTCTCAAGTTCAATTTTACCTGTGTGTCTTGAATTCGAGCAATTCCT 1219  
QY 1141 AAAAAGCATTTTGTGCAAGCCCTTGTGACATATACAGTACAGTGTCTGAAAAGCGCAGG 1200

Db 1220 AAAAAGCATTTTGTGAGCCCTTGTGACATATCCAGTACAGTGTCTAATAAGCCAGG 1279  
QY 1201 CTACAGAGCTGTGACAGTAAACAGAGGGGAGCCGATTAATAGCCGCTTTGTACAGATGCG 1260  
Db 1280 CTACAGAGCTGTGACAGTAAACAGAGGGGAGCCGATTAATAGCCGCTTTGTACAGATGCG 1339  
QY 1261 TGTGCTGCTGTGTGATCTCTCTCGCTTCCCTTCCCTTCTTTCGACAGCACACTCACTTC 1320  
Db 1340 TGTGCTGCTGTGTGATCTCTCTCGCTTCCCTTCCCTTCTTTCGACAGCACACTCACTTC 1399  
QY 1321 CTGCTGAAACATCTTCTTAACTTCAACCCAGACATATTCGTTGACAGCTCAAGTTTA 1380  
Db 1400 CTGCTGAAACATCTTCTTAACTTCAACCCAGACATATTCGTTGACAGCTCAAGTTTA 1459  
QY 1381 TTTCAACCAAGAAAGCTCCATTTTGTCTTCAACATTTGTGAAATTTGTCTACTGACACA 1440  
Db 1460 TTTCAACCAAGAAAGCTCCATTTTGTCTTCAACATTTGTGAAATTTGTCTACTGACACA 1519  
QY 1441 ACAGAGTTCTGCGGAGGAGATATGTACAGCTGTGCTGCTTGTGTGCTTCACTT 1500  
Db 1520 ACAGAGTTCTGCGGAGGAGATATGTACAGCTGTGCTGCTTGTGTGCTTCACTT 1579  
QY 1501 CTTCAGCCAAACATATGATGATCCCATGGAAGACAGGGGAAAGCCCTGGCTCTTAAGATA 1560  
Db 1580 CTTCAGCCAAACATATGATGATCCCATGGAAGACAGGGGAAAGCCCTGGCTCTTAAGATA 1639  
QY 1561 TCCATCTCTCTGCAACCAAAATTTTCACTT 1595  
Db 1640 TCCATCTCTCTGCAACCAAAATTTTCACTT 1674

RESULT 13  
ACN42470  
ID ACN42470 standard; cDNA; 3189 BP.  
XX  
AC ACN42470;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Human diagnostic and therapeutic polynucleotide SEQ ID NO:1345.  
XX  
KW seq; gene; gene therapy; human diagnostic and therapeutic polynucleotide;  
KW dltnp.  
XX  
OS Homo sapiens.  
XX  
PN W02004023973-A2.  
XX  
XX 25-MAR-2004.  
PD  
XX  
PF 12-SEP-2003; 2003MO-US028227.  
XX  
PR 12-SEP-2002; 2002US-0410259P.  
PR 12-SEP-2002; 2002US-0410260P.  
XX  
PA (INCY-) INCYTE CORP.  
XX  
PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F,  
PI Hartshorne TA, Suchorolski MT, Altus CM, Plets SJ, Elder LV;  
PI Mooney EM, Delegeane AM, Panesar IS, Barnville SC, Reddy TP;  
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstein EH;  
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LJ;  
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vilec UA, Kilton ES;  
PI Xu Y, Kwong M, Pollock JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;  
PI Pectury S, Shi X, Suarez CJ;  
XX  
XX WPI; 2004-329368/30.  
DR P-PSDB; ABM63818.  
XX  
XX  
PT New diagnostic and therapeutic polynucleotides and polypeptides, useful  
PT in diagnosing a condition, disease or disorder associated with human  
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or

PT in gene mapping.

XX Claim 1; Page; 190pp; English.

XX The invention relates to novel diagnostic and therapeutic polynucleotides  
CC selected from one of the 2722 sequences defined in the specification. A  
CC polynucleotide of the invention may have a use in gene therapy. The human  
CC diagnostic and therapeutic polynucleotides (dthp) or polypeptides may be  
CC used to diagnose a particular condition, disease or disorder associated  
CC with human molecules, e.g. cell proliferative disorder, endocrine  
CC autoimmune/inflammatory disorder, developmental disorder, endocrine  
CC disorder, neurological disorders, gastrointestinal disorders, or  
CC infections caused by virus, bacteria, fungi or parasite. The dthp  
CC molecules may also be used in genetic mapping, in identifying individuals  
CC from minute biological samples, in detecting single nucleotide  
CC polymorphisms, as molecular weight markers, and for somatic or germline  
CC gene therapy. The present sequence represents a dthp polynucleotide of  
CC the invention. Note: The sequence data for this patent is not represented  
CC in the printed specification, but was obtained in electronic format  
CC directly from WIPO at [www.wipo.int/pct/en/sequences/listing.htm](http://www.wipo.int/pct/en/sequences/listing.htm)

XX Sequence 3189 BP; 916 A; 679 C; 665 G; 929 T; 0 U; 0 Other;

Query Match 45.7%; Score 956; DB 13; Length 3189;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1056; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGAGAGGTTCTGTTACTATATGCTACACAGACAGGACGCAAGCCATCCAGAA 60  
DB 112 ATGAGAGAGGTTCTGTTACTATATGCTACACAGACAGGACGCAAGCCATCCAGAA 171  
QY 61 GAAATGTGTGAGCACTGTGTGACATGATTTTCTGAGATCTTCACTGATTAAGTAA 120  
DB 172 GAAATGTGTGAGCACTGTGTGACATGATTTTCTGAGATCTTCACTGATTAAGTAA 231  
QY 121 TCCGATATGATGACCTAAACCCGAAACAGCTCTCTGTTGTTGTTCTACACG 180  
DB 232 TCCGATATGATGACCTAAACCCGAAACAGCTCTCTGTTGTTGTTCTACACG 291  
QY 181 GGCACCGGAGACCCACCCGACACAGCCCGCAAGTTGTTAAGAAATACAAACAA 240  
DB 292 GGCACCGGAGACCCACCCGACACAGCCCGCAAGTTGTTAAGAAATACAAACAA 351  
QY 241 CTGCGGTTGATTTCTTTGCTCACTGCGGTATGGTTACTGGGTCTCGGTATTCGAA 300  
DB 352 CTGCGGTTGATTTCTTTGCTCACTGCGGTATGGTTACTGGGTCTCGGTATTCGAA 411  
QY 301 TACACCTACTTTTGCATGAGGAGGAGATTAATGATTAACGACTTCAAGAGCTTGAGCC 360  
DB 412 TACACCTACTTTTGCATGAGGAGGAGATTAATGATTAACGACTTCAAGAGCTTGAGCC 471  
QY 361 CCGCATTTCTATGACACTGACATGACATGATGATGATGATTAAGTCTTGTTGAG 420  
DB 472 CCGCATTTCTATGACACTGACATGACATGATGATGATGATTAAGTCTTGTTGAG 531  
QY 421 CCGTGAATGCTGAGATCTGCGGACGCTTCAAGAAACATTTTGTGCTCAAGAGAGCA 480  
DB 532 CCGTGAATGCTGAGATCTGCGGACGCTTCAAGAAACATTTTGTGCTCAAGAGAGCA 591  
QY 481 GAGGAGATTAAGTGGCCACTCCGCGTGCATCACTGACCTTCTTGAGGACAGACTTGTG 540  
DB 592 GAGGAGATTAAGTGGCCACTCCGCGTGCATCACTGACCTTCTTGAGGACAGACTTGTG 651  
QY 541 AAGTCAGAGCTGTACACATTAATCAAGTCAGAGCTTCAAGATTCAGATTCAGGA 600  
DB 652 AAGTCAGAGCTGTACACATTAATCAAGTCAGAGCTTCAAGATTCAGATTCAGGA 711  
QY 601 AGAAGAGATTCAGGTTTGAAGCAAAATGACGTAACGCAACCAATCCATGTTGTA 660  
DB 712 AGAAGAGATTCAGGTTTGAAGCAAAATGACGTAACGCAACCAATCCATGTTGTA 771  
QY 661 ATTGAAGACTTGAAGTCTCACTTACCCGTTCCGTTACCCCACTCTCAAGGCTCTCTG 720

DB 772 ATTGAAGACTTGAAGTCTCACTTACCCGTTCCGTTACCCCACTCTCAAGGCTCTCTG 831  
QY 721 AATATTCCTGTTTACCCCGAGATATTTTACAGTACATCTGACAGAGTCTTGCCAG 780  
DB 832 AATATTCCTGTTTACCCCGAGATATTTTACAGTACATCTGACAGAGTCTTGCCAG 891  
QY 781 GAGGAAAGCAGATATCTGTGACTTCAAGATTCAGATTTTCAAGTGCATTTCAAG 840  
DB 892 GAGGAAAGCAGATATCTGTGACTTCAAGATTCAGATTTTCAAGTGCATTTCAAG 951  
QY 841 GCAATTCACCTTACTACGATATGATGCAATTAACCACTCTGCTGTGATTAATGACAT 900  
DB 952 GCAATTCACCTTACTACGATATGATGCAATTAACCACTCTGCTGTGATTAATGACAT 1011  
QY 901 TCAATATACAGCTTTTCCATGAGCTGAGATGCTTCAAGGATGATGCTGCTTACAGT 960  
DB 1012 TCAATATACAGCTTTTCCATGAGCTGAGATGCTTCAAGGATGATGCTGCTTACAGT 1071  
QY 961 GATTCGAGGTACAAAGCTTCTCAAGATCTGACAGCTTGAAGATTAAGAGACACTGC 1020  
DB 1072 GATTCGAGGTACAAAGCTTCTCAAGATCTGACAGCTTGAAGATTAAGAGACACTGC 1131  
QY 1021 GTCTTTTGAATAATTAAGGACACACAAAGAAAGG 1058  
DB 1132 GTCTTTTGAATAATTAAGGACACACAAAGAAAGG 1169  
RESULT 14  
ADQ39029  
ID ADQ39029 standard; DNA; 3256 BP.  
XX  
AC ADQ39029;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Human SNP containing myocardial infarction-associated gene, SEQ ID 692.  
XX  
KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;  
KW cardiant; gene therapy; human; gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN MO2004058052-A2.  
XX  
PD 15-JUL-2004.  
XX  
PF 22-DEC-2003; 2003WO-US040978.  
XX  
PR 20-DEC-2002; 2002US-0434778P.  
PR 10-MAR-2003; 2003US-0451335P.  
PR 30-APR-2003; 2003US-0466412P.  
PR 23-SEP-2003; 2003US-0504955P.  
XX  
PA (APPL-) APPLERA CORP.  
XX  
PI Cargill M, Devlin JJ, Iakubova O;  
XX  
XX WPI; 2004-533949/51.  
DR P-PSDB; ADQ39857.  
XX  
XX  
PT Identifying an individual who has an altered risk for developing  
PT myocardial infarction by detecting a single nucleotide polymorphism in  
PT the individual's nucleic acids.  
XX  
PS Claim 7; SEQ ID NO 692; 145bp; English.  
XX  
CC The invention relates to a novel method for identifying an individual who  
CC has an altered risk for developing myocardial infarction. The method  
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of  
CC the nucleotide sequences given in the specification in the individual's  
CC nucleic acids, where the presence of the SNP is correlated with an  
CC altered risk for myocardial infarction in the individual. The invention  
CC further comprises: an isolated nucleic acid molecule comprising at least

CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in  
CC the specification or its complement and encoding any one of the amino  
CC acid sequences given in the specification; an isolated polypeptide  
CC comprising an amino acid sequence given in the specification; an antibody  
CC that specifically binds to the polypeptide or its antigen-binding  
CC fragment; an amplified polynucleotide containing an SNP given in the  
CC specification and which is between about 16 and 1000 nucleotides in  
CC length; a kit for detecting an SNP in a nucleic acid, comprising the  
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a  
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a  
CC method for identifying an agent useful in treating or preventing  
CC myocardial infarction. The novel detection method has cardant activity.  
CC The nucleic acids in the invention may be used in gene therapy. The  
CC method is useful in identifying an individual who has an increased or  
CC decreased risk for developing myocardial infarction and for preparing a  
CC composition for treating or preventing myocardial infarction. This  
CC polynucleotide sequence represents a human myocardial infarction-  
CC associated gene containing one or more SNPs of the invention. Note: This  
CC sequence was not shown in the specification. The sequence has come from  
CC an electronic sequence listing downloaded from the WIPO website.

XX Sequence 3256 BP; 927 A; 691 C; 669 G; 940 T; 0 U; 29 Other;

Query Match 42.8%; Score 896; DB 13; Length 3256;

Best Local Similarity 99.1%; Pred. No. 0; Mismatches 15; Indels 0; Gaps 0;

Matches 1646; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 67 TGTGAGCAAGCTGTGATCATGATTTTCTGAGATCTTCATGATTAATGATCCGAT 126  
Db 160 TGTGAGCAAGCTGTGATCATGATTTTCTGAGATCTTCATGATTAATGATCCGAT 219  
QY 127 AAGTATGACCTAAACCCGAAACAGCTCTCTGTTGTGTGTTTTCACAGGCGAC 186  
Db 220 AAGTATGACCTAAACCCGAAACAGCTCTCTGTTGTGTGTTTTCACAGGCGAC 279  
QY 187 GAGAGCCACCCGACAGCCCGCAAGTTTGAAGAAATACAGAACCAACATCCG 246  
Db 280 GAGAGCCACCCGACAGCCCGCAAGTTTGAAGAAATACAGAACCAACATCCG 339  
QY 247 GTTGATTTCTTGTCTCACTGCGGTATGGTTACTGCGTCTCGGTATTCAGATAC 306  
Db 340 GTTGATTTCTTGTCTCACTGCGGTATGGTTACTGCGTCTCGGTATTCAGATAC 399  
QY 307 TACTTTGCAATGGGGGGAAGATTAATGATTAACGACTTCAAGGCTTGAAGCCG 366  
Db 400 TACTTTGCAATGGGGGGAAGATTAATGATTAACGACTTCAAGGCTTGAAGCCG 459  
QY 367 TTCTATGACCTGACATGACATGATGATGTTAGAACTTGTGTGAGCCGTG 426  
Db 460 TTCTATGACCTGACATGACATGATGATGTTAGAACTTGTGTGAGCCGTG 519  
QY 427 ATTGCTGACCTGCGCAGCCCTCAAGAACATTTTGTAGCAAGCAGAGCAAGAG 486  
Db 520 ATTGCTGACCTGCGCAGCCCTCAAGAACATTTTGTAGCAAGCAGAGCAAGAG 579  
QY 487 ATAAGTGGCGCATCCCGGTGGCATCACTGCACTCTTGAAGACAGACCTTGAAG 546  
Db 580 ATAAGTGGCGCATCCCGGTGGCATCACTGCACTCTTGAAGACAGACCTTGAAG 639  
QY 547 GAGCTGTACACATGATCAAGTGAAGTGTGATTCAGATTCAGGAAGAA 606  
Db 640 GAGCTGTACACATGATCAAGTGAAGTGTGATTCAGATTCAGGAAGAA 699  
QY 607 GATTTGAGGTTTGAAGCAAAATGACAGTAAACAGCAACCAATCTGATTAAT 666  
Db 700 GATTTGAGGTTTGAAGCAAAATGACAGTAAACAGCAACCAATCTGATTAAT 759  
QY 667 GACTTGAATCTCACTTACCGTTGCGTACCCCACTTCAAGCCTCTGATAT 726  
Db 760 GACTTGAATCTCACTTACCGTTGCGTACCCCACTTCAAGCCTCTGATAT 819  
QY 727 CTTGGTTTACCCCGAATATTTTACAGTACATGAGAGATCTTGGCCAGAGAA 786

Db 820 CTTGGTTTACCCCGAATATTTTACAGTACATGAGAGATCTTGGCCAGAGAA 879  
QY 787 AGCAAGATCTGTGACCTTCAAGACATCCAGTTTTCAGAGTCCAAATTTCAAGGAG 846  
Db 880 AGCAAGATCTGTGACCTTCAAGACATCCAGTTTTCAGAGTCCAAATTTCAAGGAG 939  
QY 847 CAATTAATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 906  
Db 940 CAATTAATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 999  
QY 907 ACAGATTTTCTATGAGCTGAGATGAGTCTTCAAGGATGATGATGATGATGATGAT 966  
Db 1000 ACAGATTTTCTATGAGCTGAGATGAGTCTTCAAGGATGATGATGATGATGATGAT 1059  
QY 967 GAGTGAACAGCTTCAAGATGAGTCTTCAAGGATGAGTCTTCAAGGATGAGTCT 1026  
Db 1060 GAGTGAACAGCTTCAAGATGAGTCTTCAAGGATGAGTCTTCAAGGATGAGTCT 1119  
QY 1027 TGAAGATTAAGGAGACCAAGAAAGAGAGGATCTTCAAGGATGATGATGATGAT 1086  
Db 1120 TGAAGATTAAGGAGACCAAGAAAGAGAGGATCTTCAAGGATGATGATGATGAT 1179  
QY 1087 GATGTTCTCTCAAGTTCATTTTACCTGATGATGATGATGATGATGATGATGAT 1146  
Db 1180 GATGTTCTCTCAAGTTCATTTTACCTGATGATGATGATGATGATGATGATGAT 1239  
QY 1147 GCATTTTTCAGAGCCCTTGTGACATTAACAGTACAGTGTGAAAGGCGAGCTAC 1206  
Db 1240 GCATTTTTCAGAGCCCTTGTGACATTAACAGTACAGTGTGAAAGGCGAGCTAC 1299  
QY 1207 GAGCTGTGAGTAAACAAAGGAGAGGAGTAAAGGAGTAAAGGAGAGAGAGAG 1266  
Db 1300 GAGCTGTGAGTAAACAAAGGAGAGGAGTAAAGGAGTAAAGGAGAGAGAGAG 1359  
QY 1267 TGTGTTGATGATCTCTCTGCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 1326  
Db 1360 TGTGTTGATGATCTCTCTGCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 1419  
QY 1327 GAAATCTTCTTAACTTCAACCCAGACCAATTTGTGTGACAGTCAAGTTATTC 1386  
Db 1420 GAAATCTTCTTAACTTCAACCCAGACCAATTTGTGTGACAGTCAAGTTATTC 1479  
QY 1387 CAGGAAAGCTCATTTTGTCTTCAATTTGTGGAATTTCTGTCTACCTGCAACA 1446  
Db 1480 CAGGAAAGCTCATTTTGTCTTCAATTTGTGGAATTTCTGTCTACCTGCAACA 1539  
QY 1447 GTTCTGGAAGGATATGATGATGATGATGATGATGATGATGATGATGATGAT 1506  
Db 1540 GTTCTGGAAGGATATGATGATGATGATGATGATGATGATGATGATGATGAT 1599  
QY 1507 CCAACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1566  
Db 1600 CCAACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1659  
QY 1567 TCTCTCTGGAACCAATTTCTTCACTTCACTTCACTTCACTTCACTTCACTTCA 1626  
Db 1660 TCTCTCTGGAACCAATTTCTTCACTTCACTTCACTTCACTTCACTTCACTTCA 1719  
QY 1627 GTGAGTCCAGAAACCGGATAGCCGTTTATTTGATGATGATGATGATGATGAT 1686  
Db 1720 GTGAGTCCAGAAACCGGATAGCCGTTTATTTGATGATGATGATGATGATGAT 1779  
QY 1687 CAGGAAACCAACCCAGATGAAATTTTGAAGCAATGTGTT 1727  
Db 1780 CAGGAAACCAACCCAGATGAAATTTTGAAGCAATGTGTT 1820

RESULT 15  
ADQ39030  
ID ADQ39030 standard; DNA; 3274 BP.  
XX  
XX ADQ39030;  
XX



DT	18-NOV-2004	(first entry)
XX	Human SNP containing myocardial infarction-associated gene, SEQ ID 693.	
XX	Myocardial infarction; detection; single nucleotide polymorphism; SNP;	
KM	cardiant; gene therapy; human; gene; ds.	
OS	Homo sapiens.	
XX	WO2004058052-A2.	
PN	15-JUL-2004.	
PD	22-DEC-2003; 2003WO-US040978.	
PF	20-DEC-2002; 2002US-0434778P.	
XX	10-MAR-2003; 2003US-0453135P.	
PR	30-APR-2003; 2003US-0466412P.	
XX	23-SEP-2003; 2003US-0504955P.	
PA	(APPL-) APPLEBA CORP.	
PI	Cargill M, Devlin J, Takolova O;	
XX	WPI; 2004-533949/51.	
DR	P-PSDB; ADQ39858.	
XX	Identifying an individual who has an altered risk for developing	
PT	myocardial infarction by detecting a single nucleotide polymorphism in	
PT	the individual's nucleic acids.	
XX	Claim 7; SEQ ID NO 693; 145pp; English.	
XX	The invention relates to a novel method for identifying an individual who	
CC	has an altered risk for developing myocardial infarction. The method	
CC	comprises detecting a single nucleotide polymorphism (SNP) in any one of	
CC	the nucleotide sequences given in the specification in the individual's	
CC	nucleic acids, where the presence of the SNP is correlated with an	
CC	altered risk for myocardial infarction in the individual. The invention	
CC	further comprises: an isolated nucleic acid molecule comprising at least	
CC	8 contiguous nucleotides where one of the nucleotides is an SNP given in	
CC	the specification or its complement and encoding any one of the amino	
CC	acid sequences given in the specification; an isolated polypeptide	
CC	comprising an amino acid sequence given in the specification; an antibody	
CC	that specifically binds to the polypeptide or its antigen-binding	
CC	fragment; an amplified polynucleotide containing an SNP given in the	
CC	specification and which is between about 16 and 1000 nucleotides in	
CC	length; a kit for detecting an SNP in a nucleic acid, comprising the	
CC	polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a	
CC	nucleic acid molecule; a method of detecting a variant polypeptide; and a	
CC	method for identifying an agent useful in treating or preventing	
CC	myocardial infarction. The novel detection method has cardiant activity.	
CC	The nucleic acids of the invention may be used in gene therapy. The	
CC	method is useful in identifying an individual who has an increased or	
CC	decreased risk for developing myocardial infarction and for preparing a	
CC	composition for treating or preventing myocardial infarction. This	
CC	polynucleotide sequence represents a human myocardial infarction-	
CC	associated gene containing one or more SNPs of the invention. Note: This	
CC	sequence was not shown in the specification. The sequence has come from	
CC	an electronic sequence listing downloaded from the WIPO website.	
XX	Sequence 3274 BP; 932 A; 694 C; 672 G; 946 T; 0 U; 30 Other;	
SO		
Query Match	42.8%; Score 896; DB 13; Length 3274;	
Best Local Similarity	99.1%; Pred. No. 0; Mismatches 15; Indels 0; Gaps 0;	
Matches 1646; Conservative	0; Mismatches 15; Indels 0; Gaps 0;	
67	TGTGAGCAAGCTGTGTGATCATGATTTCTGCGAGATCTTCACTGATTAAGTGAATCCGAT 126	
Db	178 TGTGAGCAAGCTGTGTGATCATGATTTCTGCGAGATCTTCACTGATTAAGTGAATCCGAT 237	
67	127 AAGTGAAGCTTAAACCGAAACAGCTCCTCTGTGTGTGTGTTCTTCAACGCGGCAAC 186	

Db	238	AAAGTATGACCTAATAAAACCGAAACAGCTCCTCTTGTGTGTGCTTCTACACAGGGACCC	297
Qy	187	GGAGACCCACCCGACACAGCCCGCAGTTGTTAAGAAATACGAACCAACACCTGCGC	246
Db	298	GGAGACCCACCCGACACAGCCCGCAGTTGTTAAGAAATACGAACCAACACCTGCGC	357
Qy	247	GTTGATTTCTTTGGTCACTGCGGTTAATGGGTTACTGGGTCGCGGTAATTGAAATPACCC	306
Db	358	GTTGATTTCTTTGGTCACTGCGGTTAATGGGTTACTGGGTCGCGGTAATTGAAATPACCC	417
Qy	307	TACTTTTGCATATGGGGGGAAGAATAATTGATPAAACGATTCAGAGCTTGGACCCCGCAT	366
Db	418	TACTTTTGCATATGGGGGGAAGAATAATTGATPAAACGATTCAGAGCTTGGACCCCGCAT	477
Qy	367	TTCTATGACACTGAGACATGACAGATGACTGTGTAGTTTGAACCTTGTGTAGCCGTGG	426
Db	478	TTCTATGACACTGAGACATGAGATGACTGTGTAGTTTGAACCTTGTGTAGCCGTGG	537
Qy	427	ATTGCTGGACTCTGGCCAGCCCTCAGAAAGCATTTTGTAGTCAACACAGAGCAAGAGAG	486
Db	538	ATTGCTGGACTCTGGCCAGCCCTCAGAAAGCATTTTGTAGTCAACACAGAGCAAGAGAG	597
Qy	487	ATTAAGTGCGCACTCCCGGTGACATCACTGTCATCTTGAAGACAGACCTTGTGAAGTCA	546
Db	598	ATTAAGTGCGCACTCCCGGTGACATCACTGTCATCTTGAAGACAGACCTTGTGAAGTCA	657
Qy	547	GAGCTGCTACACATTGAACTTCAAGTGAAGCTTGTGAATTCAGATTCAGAAAGAA	606
Db	658	GAGCTGCTACACATTGAACTTCAAGTGAAGCTTGTGAATTCAGATTCAGAAAGAA	717
Qy	607	GATTCTGAGTTTGAAGCAAAATGCAGTGAACAGCAACCAATCCATGTGTGAATTGAA	666
Db	718	GATTCTGAGTTTGAAGCAAAATGCAGTGAACAGCAACCAATCCATGTGTGAATTGAA	777
Qy	667	GACTTTGAGTCTGACTTAACCGTGGGTAACCCCACTGCAAGAGCTCTGAAATATT	726
Db	778	GACTTTGAGTCTGACTTAACCGTGGGTAACCCCACTGCAAGAGCTCTCTGAAATATT	837
Qy	727	CTGTGTTAACCCCGAGAAATATTACAGGTACATCTGACAGAGAGTCTTGTGGCAGAGGAA	786
Db	838	CTGTGTTAACCCCGAGAAATATTACAGGTACATCTGACAGAGAGTCTTGTGGCAGAGGAA	897
Qy	787	AGCCAAGTATCTGTGACTTCACAGCAGATCCAGTTTTCAGATGCCAATTTCAAAGCAGTT	846
Db	898	AGCCAAGTATCTGTGACTTCACAGCAGATCCAGTTTTCAGATGCCAATTTCAAAGCAGTT	957
Qy	847	CAACTTACTAGATGATGCAATGCAATPAAACCACTGCGTGGTAAGAAATTGGAATTTCAAT	906
Db	958	CAACTTACTAGATGATGCAATGCAATPAAACCACTGCGTGGTAAGAAATTGGAATTTCAAT	1017
Qy	907	ACAGACTTTTCTATACAGCTGAGAGATGCTTCAGCGTGAATCTGCCCTPACAGTGAATCT	966
Db	1018	ACAGACTTTTCTATACAGCTGAGAGATGCTTCAGCGTGAATCTGCCCTPACAGTGAATCT	1077
Qy	967	GAGGTACAAAGCCTACTCCAAAGACTGCAAGCTTGAAGATPAAAGAGACACTGCGTCTT	1028
Db	1078	GAGGTACAAAGCCTACTCCAAAGCTGCAAGCTTGAAGATPAAAGAGACACTGCGTCTT	1137
Qy	1027	TTTGAATAAAGGAGAGACACAAATAAGAAAGAGACTACCTTAACCCGACATATACCTGG	1088
Db	1138	TTTGAATAAAGGAGAGACACARBAAGAAAGAGACTACCTTAACCCGACATATACCTGG	1197
Qy	1087	GGATGTTCTCTCCAGTTCAATTTTACCTGTGTCTTGAATCCGAGCAATTCCTPAAAG	1146
Db	1198	GGATGTTCTCTCCAGTTCAATTTTACCTGTGTCTTGAATCCGAGCAATTCCTPAAAG	1255
Qy	1147	GCATTTTTCGAGCCCTTGTGAATPACAGTGAACAGTGTCAAAAGCCGACGCTACAG	1206
Db	1258	GCATTTTTCGAGCCCTTGTGAATPACAGTGAACAGTGTCAAAAGCCGACGCTACAG	1317
Qy	1207	GAGCTGTGCAATAAACAAGGGGAGCCGATTAATAGCCGCTTTGTACAGAGTCCGTGACC	1266
Db	1318	GAGCTGTGCAATAAACAAGGGGAGCCGATTAATAGCTGCTTTGTACAGAGTCCGTGACC	1377



```
QY 1267 TGCCTTGGATCTCCTCTGCTTCCCTTTCCTTTCGAGCCACCACTCAGTCTCTGCTC 1326
    |||||
Db 1378 TGCCTTGGATCTCCTCCTCGCTTTCCTTTCGAGCCACCACTCAGTCTCTGCTC 1437
    |||||
QY 1327 GAACATCTTCTTAACTTCAACCCAGACCATTTCTGTGCAAGCTCAAGTTATTTCAC 1386
    |||||
Db 1438 GAACATCTTCTTAACTTCAACCCAGACCATTTCTGTGCAAGCTCAAGTTATTTCAC 1497
    |||||
QY 1387 CCAGAAAGCTCCATTTTGTCTTCAACATTTGTGGAATTTCTGTACTGCCACAAGAG 1446
    |||||
Db 1498 CCAGAAAGCTCCATTTTGTCTTCAACATTTGTGGAATTTCTGTACTGCCACAAGAG 1557
    |||||
QY 1447 GTTCTGCGAAGGAGATGATGACAGGCTGGCTGCTTGTGTGCTTCAAGTTCTTCAG 1506
    |||||
Db 1558 GTTCTGCGAAGGAGATGATGACAGGCTGGCTGCTTGTGTGCTTCAAGTTCTTCAG 1617
    |||||
QY 1507 CCAACATTAATGATCCCATGGAAGACGCGGAAAGCCCTGGCTCCTAAGATATCCATC 1566
    |||||
Db 1618 CCAACATTAATGATCCCATGGAAGACGCGGAAAGCCCTGGCTCCTAAGATATCCATC 1677
    |||||
QY 1567 TCTCTGGAACAACAATTTCTTCCACTTACAGATGACCCCTCAATCCCATCATPATG 1626
    |||||
Db 1678 TCTCTGGAACAACAATTTCTTCCACTTACAGATGACCCCTCAATCCCATCATPATG 1737
    |||||
QY 1627 GTGGGTCCAGAAACCGGCAATAGCCCGTTTATTGGGTTCTCTACAACATAGAGAAATC 1686
    |||||
Db 1738 GTGGGTCCAGAAACCGGCAATAGCCCGTTTATTGGGTTCTCTACAACATAGAGAAATC 1797
    |||||
QY 1687 CAAGAACACACCCAGATGGAATTTTGAGCAATGTGTT 1727
    |||||
Db 1798 CAAGAACACACCCAGATGGAATTTTGAGCAATGTGTT 1838
    |||||
```

Search completed: August 27, 2005, 01:18:41  
Job time : 737.44 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 23:30:25 ; Search time 235.42 Seconds  
(without alignments)  
14554.251 Million cell updates/sec

Title: US-09-371-347a-45

Perfect score: 2094

Sequence: 1 atgagagaggttcgttact.....ttcagatattgtgcataa 2094

Scoring table: OLIGO\_NUC

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/prodata/1/ina/5B\_COMB.seq:\*
- 2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq:\*
- 3: /cgn2\_6/prodata/1/ina/5A\_COMB.seq:\*
- 4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*
- 5: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*
- 6: /cgn2\_6/prodata/1/ina/Backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1854	88.5	3259	US-09-318-448-23	Sequence 23, Appl
2	1701	81.2	3242	US-09-949-016-4215	Sequence 4215, Ap
3	386	18.4	390	US-08-905-223-71	Sequence 71, Appl
4	330	15.8	601	US-09-949-016-150019	Sequence 150019,
5	330	15.8	35916	US-09-949-016-15957	Sequence 15957, A
6	279	13.3	601	US-09-949-016-150020	Sequence 150020,
7	189	9.0	601	US-09-949-016-150037	Sequence 150037,
8	158	7.5	2475	US-09-566-921-88	Sequence 88, Appl
9	155	7.4	601	US-09-949-016-150030	Sequence 150030,
10	145	6.9	601	US-09-949-016-150031	Sequence 150031,
11	137	6.5	601	US-09-949-016-150046	Sequence 150046,
12	137	6.5	601	US-09-949-016-150047	Sequence 150047,
13	125	6.0	601	US-09-949-016-150029	Sequence 150029,
14	121	5.8	601	US-09-949-016-150041	Sequence 150041,
15	121	5.8	601	US-09-949-016-150042	Sequence 150042,
16	119	5.7	601	US-09-949-016-150008	Sequence 150008,
17	119	5.7	601	US-09-949-016-150055	Sequence 150055,
18	110	5.3	601	US-09-949-016-150048	Sequence 150048,
19	94	4.5	601	US-09-949-016-150032	Sequence 150032,
20	78	3.7	244	US-09-471-276-495	Sequence 495, App
21	78	3.7	601	US-09-949-016-150007	Sequence 150007,
22	76	3.6	601	US-09-949-016-150018	Sequence 150018,
23	30	1.4	1681	US-09-023-655-453	Sequence 453, Appl
24	20	1.0	273	US-09-513-999C-14761	Sequence 14761, A
25	20	1.0	440	US-09-397-787-305	Sequence 305, Appl
26	20	1.0	444	US-09-621-976-14139	Sequence 14139, A
27	20	1.0	445	US-09-397-787-274	Sequence 274, App

C 28	20	1.0	174259	4	US-09-949-016-11968	Sequence 11968, A
C 29	20	1.0	174262	4	US-09-949-016-14259	Sequence 14259, A
C 30	19	0.9	169	1	US-08-166-346A-8	Sequence 8, Appl
C 31	19	0.9	459	4	US-09-621-976-8324	Sequence 8324, Ap
C 32	19	0.9	3969	3	US-09-518-386B-4	Sequence 4, Appl
C 33	19	0.9	4396	3	US-09-821-736-1	Sequence 1, Appl
C 34	19	0.9	14721	4	US-09-949-016-13507	Sequence 13507, A
C 35	19	0.9	25199	4	US-09-949-016-13361	Sequence 13361, A
C 36	19	0.9	129658	4	US-09-949-016-17195	Sequence 17195, A
C 37	19	0.9	186734	4	US-09-949-016-14870	Sequence 14870, A
C 38	19	0.9	193689	4	US-09-949-016-12350	Sequence 12350, A
C 39	19	0.9	193689	4	US-09-949-016-13088	Sequence 13088, A
C 40	19	0.9	200663	4	US-09-949-016-12569	Sequence 12569, A
C 41	19	0.9	203093	4	US-09-949-016-14445	Sequence 14445, A
C 42	18	0.9	78	2	US-08-749-852-56	Sequence 56, Appl
C 43	18	0.9	78	2	US-08-749-852-58	Sequence 58, Appl
C 44	18	0.9	511	4	US-09-902-540-1374	Sequence 1374, Ap
C 45	18	0.9	531	4	US-09-252-991A-2223	Sequence 2223, Ap

#### ALIGNMENTS

RESULT 1						
US-09-318-448-23						
Sequence 23, Application US/09318448						
Patent No. 6210950						
GENERAL INFORMATION:						
APPLICANT: Johnson, William G.						
APPLICANT: Steenroos, Edward S.						
TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING						
TITLE OF INVENTION: DEVELOPMENTAL DISORDERS						
FILE REFERENCE: 601-1-057						
CURRENT APPLICATION NUMBER: US/09/318,448						
CURRENT FILING DATE: 1999-05-25						
NUMBER OF SEQ ID NOS: 46						
SOFTWARE: PatentIn Ver. 2.0						
SEQ ID NO 23						
LENGTH: 3259						
TYPE: DNA						
ORGANISM: Homo sapiens						
US-09-318-448-23						
Query Match						
Best Local Similarity 99.9%; Pred. No. 0;						
Matches 2094; Conservative 0; Mismatches 0; Indels 3; Gaps 1;						
QY	1	ATGAGAGGTTCTGTACTATATGCTACACAGGAGGAGCAAGGCCATCGCAGAA	60			
DB	80	ATGAGAGGTTCTGTACTATATGCTACACAGGAGGAGCAAGGCCATCGCAGAA	139			
QY	61	GAATGTGTGACCAAGCTGTGTACATGATTTTGTGAGATCTTCACTGTATAGTAA	120			
DB	140	GAATGTGTGACCAAGCTGTGTACATGATTTTGTGAGATCTTCACTGTATAGTAA	199			
QY	121	TCCGATAGTATGACTTAAACCCGAAACGCTCTTGTGTGTGTCTTCTCCAG	180			
DB	200	TCCGATAGTATGACTTAAACCCGAAACGCTCTTGTGTGTGTCTTCTCCAG	259			
QY	181	GGCAGCGAGAGCCACCCGACAGCCCGCAGTTGTTAAGAAATACAGAACCAACA	240			
DB	260	GGCAGCGAGAGCCACCCGACAGCCCGCAGTTGTTAAGAAATACAGAACCAACA	319			
QY	241	CTGCCGTTGATTTCTTCTCACTGCGGTATGGTTACTGGTCTCGGTATTCAGAA	300			
DB	320	CTGCCGTTGATTTCTTCTCACTGCGGTATGGTTACTGGTCTCGGTATTCAGAA	379			
QY	301	TACACTTCTTGTGAATGGGGGAAATATGATTAACGACTTCAAGAGCTGGAGCC	360			
DB	380	TACACTTCTTGTGAATGGGGGAAATATGATTAACGACTTCAAGAGCTGGAGCC	439			
QY	361	CGGATTTCTATGACACTGACATGACATGACTGTGATGATTTGATTTGATTTGAG	420			

Db	440	GGGCAATTTCTATGACACCTGACCAATGCAAGTACATGCTGTAGAGTTTAGAACTTTGGTTGAG	499
Qy	421	CCGTGGAATGCTGGACCTCTGCGCAGCCCTCAGAAAGCAATTTTAGTCAAGCAGAGCA	480
Db	500	CCGTGGAATGCTGGACCTCTGCGCAGCCCTCAGAAAGCAATTTTAGTCAAGCAGAGCA	559
Qy	481	GAGGAGATTAAGTGGCGCAGCTCCGGGTGGCACTCATGATCTCTTGAGGACAGCACTTGTG	540
Db	560	GAGGAGATTAAGTGGCGCAGCTCCGGGTGGCACTCATGATCTCTTGAGGACAGCACTTGTG	619
Qy	541	AAGTCAGAGCTGCTACACATTTGAATCTCAAGTCAGCTTCTGAGATTCGATATTCAGGA	600
Db	620	AAGTCAGAGCTGCTACACATTTGAATCTCAAGTCAGCTTCTGAGATTCGATATTCAGGA	679
Qy	601	AGAAAGATTTCTGAGGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATCCATGTTGTGA	660
Db	680	AGAAAGATTTCTGAGGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATCCATGTTGTGA	739
Qy	661	ATTGAAGCTTTGAGTCTCTCACTTAACCGGTGGTAACCCGTCATCAAGCCCTCTG	720
Db	740	ATTGAAGCTTTGAGTCTCTCACTTAACCGGTGGTAACCCGTCATCAAGCCCTCTCTG	799
Qy	721	AATATTCCTGGTTTACCCCGAGAAATATTTACAGGTACATCTGCAGAGTCTCTTGGCAG	780
Db	800	AATATTCCTGGTTTACCCCGAGAAATATTTACAGGTACATCTGCAGAGTCTCTTGGCAG	859
Qy	781	GAGAAAGCCAAATCTGTGACTTCAGCAGATCCAGTTTTCAGATGCCAATTTCAAAG	840
Db	860	GAGAAAGCCAAATCTGTGACTTCAGCAGATCCAGTTTTCAGATGCCAATTTCAAAG	919
Qy	841	GCAGTTCAACTTACTAGCAATGATGCCATAAAAACACTGCTGCTGTGAATTTGACATTT	900
Db	920	GCAGTTCAACTTACTAGCAATGATGCCATAAAAACACTGCTGCTGTGAATTTGACATTT	979
Qy	901	TCAATATCAGACTTTTCCCTATCAGGCTGGAATGCTTCAGGCTGATCTGCTTAAAGT	960
Db	980	TCAATATCAGACTTTTCCCTATCAGGCTGGAATGCTTCAGGCTGATCTGCTTAAAGT	1039
Qy	961	GATTCTGAGTACAAAGCTTACTCCAAAGACTGACGACTTGAAGATAAAGAGACACTGC	1020
Db	1040	GATTCTGAGTACAAAGCTTACTCCAAAGACTGACGACTTGAAGATAAAGAGACACTGC	1099
Qy	1021	GTCTTTTGAATAATTAAGGAGCAGACAAAGAAAGAGAGCTTACCTTACCCACATATA	1080
Db	1100	GTCTTTTGAATAATTAAGGAGCAGACAAAGAAAGAGAGCTTACCTTACCCACATATA	1159
Qy	1081	CCTGGGGAGTGTCTCTCAGTTCATTTTACCTGCTGCTTGAATCCAGCAATTCCT	1140
Db	1160	CCTGGGGAGTGTCTCTCAGTTCATTTTACCTGCTGCTTGAATCCAGCAATTCCT	1219
Qy	1141	AAAAAGGCAATTTTGTGCAAGCCCTGTGTGACTATACCAAGTGAACAGTCTGAAGACGCAAG	1200
Db	1220	AAAAAGGCAATTTTGTGCAAGCCCTGTGTGACTATACCAAGTGAACAGTCTGAAGACGCAAG	1279
Qy	1201	CTACAGAGCTGTGCAATTAACAAAGGGGCAAGCCGATTATAGCCGTTTGTACAGATGCC	1260
Db	1280	CTACAGAGCTGTGCAATTAACAAAGGGGCAAGCCGATTATAGCCGTTTGTACAGATGCC	1339
Qy	1261	TGTGCTGCTGTTGGATCTCTCTCTGCTTCCCTTCTGCAAGCCCACTAGTCTC	1320
Db	1340	TGTGCTGCTGTTGGATCTCTCTCTGCTTCCCTTCTGCAAGCCCACTAGTCTC	1399
Qy	1321	CTGCTCGAATCTTCTCTTAACTTCAACCCAGACCATATTCGTGTGCAAGCTCAAGTTTA	1380
Db	1400	CTGCTCGAATCTTCTCTTAACTTCAACCCAGACCATATTCGTGTGCAAGCTCAAGTTTA	1459
Qy	1381	TTTCAACCCAGAAAGCTTCATTTTGTCTTCAACATTTGTGAATTTCTGTCTACTGCCACA	1440
Db	1460	TTTCAACCCAGAAAGCTTCATTTTGTCTTCAACATTTGTGAATTTCTGTCTACTGCCACA	1519
Qy	1441	ACAGAGGTTTCCCGAAGAGGAGTATGTACAGGCTGGCGGCTTGTGGTTCTTCACTT	1500
Db	1520	ACAGAGGTTTCCCGAAGAGGAGTATGTACAGGCTGGCGGCTTGTGGTTCTTCACTT	1579

[illegible]

80 ATGAGAGGTTTCTGTACTATATGCTACACAGAGGACAGGCAAGGCCATCCGACAA 139  
61 GAAATGTGACAGACGTGTGTACATGGAATTTCTGAGATCTTCACTGATTTAGTAA 120  
140 GAAATATGACAGACGTGTGTACATGGAATTTCTGAGATCTTCACTGATTTAGTAA 199  
121 TCCGATTAAGTATGACCTTAATAAACCCGAAACAGCTCTCTGTTGTGTGTTCTACACG 180  
200 TCCGATTAAGTATGACCTTAATAAACCCGAAACAGCTCTCTGTTGTGTGTTCTACACG 259  
181 GGCACCGAGAGACCCACCCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACA 240  
260 GGCACCGAGAGACCCACCCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACA 319  
241 CTGCGCGTGAATTTCTTTGCTCACTGCGGTATGAGTTATCTGGGTCTCGGTATTCAGAA 300  
320 CTGCGCGTGAATTTCTTTGCTCACTGCGGTATGAGTTATCTGGGTCTCGGTATTCAGAA 379  
301 TACACTTACTTTTGCAATGGGGGAAAGATTAATTAACAAGCTTCAAGAGCTTGAGCC 360  
380 TACACTTACTTTTGCAATGGGGGAAAGATTAATTAACAAGCTTCAAGAGCTTGAGCC 439  
361 GCGCATTTCTATGACATGAGACATGAGATGACTGTAGGTTTAAAGCTTGTGTAG 420  
440 GCGCATTTCTATGACATGAGACATGAGATGACTGTAGGTTTAAAGCTTGTGTAG 499  
421 CCGTGAATGCTGAGATCTGCGGACGCTTCAGAAAGATTTTATAGTCAAGAGAGACAA 480  
500 CCGTGAATGCTGAGATCTGCGGACGCTTCAGAAAGATTTTATAGTCAAGAGAGACAA 559  
481 GAGAGATTAAGTGGGACATCCCGGTGGATCACTGATCCCTTGAGAGACAGCTTGTG 540  
560 GAGAGATTAAGTGGGACATCCCGGTGGATCACTGATCCCTTGAGAGACAGCTTGTG 619  
541 AAGTCAGAGCTGCTACATGATTAATCTCAAGTCAGCTTCTGAGATTCAGATTCAGGA 600  
620 AAGTCAGAGCTGCTACATGATTAATCTCAAGTCAGCTTCTGAGATTCAGATTCAGGA 679  
601 AGAAGAGATTTGAGGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATCCATGTTGTA 660  
680 AGAAGAGATTTGAGGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATCCATGTTGTA 739  
661 ATTGAAGCTTTGAGGCTCACTTACCCGTTGGTACCCCACTCTCAAGGCTCTCTG 720  
740 ATTGAAGCTTTGAGGCTCACTTACCCGTTGGTACCCCACTCTCAAGGCTCTCTG 799  
721 AATATTTCTGTGTTTACCCCAAGATATTTACAGTATCATCTGACAGAGTCTCTTGCCAG 780  
800 AATATTTCTGTGTTTACCCCAAGATATTTACAGTATCATCTGACAGAGTCTCTTGCCAG 859  
781 GAGGAAAGCCAGATATCTGTGACTTCAAGATCCAGTTTTCAGTCCCAATTTCAAG 840  
860 GAGGAAAGCCAGATATCTGTGACTTCAAGATCCAGTTTTCAGTCCCAATTTCAAG 919  
841 GAGGATCACTTACTGAGATGATGCAATTAACCACTGCTGTGTAGATTGACATT 900  
920 GAGGATCACTTACTGAGATGATGCAATTAACCACTGCTGTGTGTAGATTGACATT 979  
901 TCAATATCAGACTTTTCTATCAGCTTGAGATGCTTCAAGCTGATCTGCCCTTAACAGT 960  
980 TCAATATCAGACTTTTCTATCAGCTTGAGATGCTTCAAGCTGATCTGCCCTTAACAGT 1039  
961 GATTCGAGATCAAAAGCTTACTCAAAAGCTGCAAGCTTGAAGATTAAGAGAGCACTGC 1020  
1040 GATTCGAGATCAAAAGCTTACTCAAAAGCTGCAAGCTTGAAGATTAAGAGAGCACTGC 1099  
1021 GTTCCTTTGAAATTAAGGACACAAAGAAAGAGCTTACTTACCCGACATATA 1080  
1100 GTTCCTTTGAAATTAAGGACACAAAGAAAGAGCTTACTTACCCGACATATA 1159  
1081 CCGTGGGAGATGTTCTCTCAAGTTCATTTTACCTGTGTCTTGAATTCGAGCAATTCCT 1140  
1160 CCGTGGGAGATGTTCTCTCAAGTTCATTTTACCTGTGTCTTGAATTCGAGCAATTCCT 1219

1141 AAAAGCATTTTTCGAGCCCTTGTGACTATACAGTGAACAGTGTGAAAAGCCAGG 1200  
1220 AAAAGCATTTTTCGAGCCCTTGTGACTATACAGTGAACAGTGTGAAAAGCCAGG 1279  
1201 CTACAGAGCTGTGTGATTAACAAAGGAGCAGCGATTAATAGCCGCTTTGTACAGATGCC 1260  
1280 CTACAGAGCTGTGTGATTAACAAAGGAGCAGCGATTAATAGCCGCTTTGTACAGATGCC 1339  
1261 TGTGCTGCTGTGTGATCTCTCTGCTTCTTCCCTTCTGACAGCCACCTCAAGTCTC 1320  
1340 TGTGCTGCTGTGTGATCTCTCTGCTTCTTCCCTTCTGACAGCCACCTCAAGTCTC 1399  
1321 CTGCTGAAACATCTTCTTAACCTTCAACCCAGACCATATGCTGTGACAGCTCAAGTTTA 1380  
1400 CTGCTGAAACATCTTCTTAACCTTCAACCCAGACCATATGCTGTGACAGCTCAAGTTTA 1459  
1381 TTTCAACCGAAGAACTCCATTTTGTCTTCAACATTTGTGAAATTTCTGTCTACCTGACA 1440  
1460 TTTCAACCGAAGAACTCCATTTTGTCTTCAACATTTGTGAAATTTCTGTCTACCTGACA 1519  
1441 ACAGAGTTCTGTGAGAGAGATATGTACAGCTGTGCTGTGTGTGTGTGTGTGTGTGTGT 1500  
1520 ACAGAGTTCTGTGAGAGAGATATGTACAGCTGTGCTGTGTGTGTGTGTGTGTGTGTGT 1579  
1501 CTTGAGCCAAACATCATGATCCCATGAAAGACAGCGGAAAGCCCTGCTCTTAAGATA 1560  
1580 CTTGAGCCAAACATCATGATCCCATGAAAGACAGCGGAAAGCCCTGCTCTTAAGATA 1639  
1561 TCCATCTCTCTGTGAAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
1640 TCCATCTCTCTGTGAAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1699  
1621 ATATATGTGTGTGTGAGAACCCGCAATAGCCGTTTATTTGGTCTCTACAAATAGAGAG 1680  
1700 ATATATGTGTGTGTGAGAACCCGCAATAGCCGTTTATTTGGTCTCTACAAATAGAGAG 1759  
1681 AAATCCAAAGAACCAACCCAGATGAAATTTTGGAGCAATGTG--GTTTTTGGCTGC 1737  
1760 AAATCCAAAGAACCAACCCAGATGAAATTTTGGAGCAATGTGTTTGTGTGTGTGTGTGT 1819  
1738 AGCATTAAGATAGGATTAATCTATTCAGAAAAGAGCTCAGACATTTCTTAAGCATGG 1797  
1820 AGCATTAAGATAGGATTAATCTATTCAGAAAAGAGCTCAGACATTTCTTAAGCATGG 1879  
1798 ATCTTAATCTATTAAGTTTCTTCTCAAGATGCTCTGTTGGGAGAGAGAGCC 1857  
1880 ATCTTAATCTATTAAGTTTCTTCTCAAGATGCTCTGTTGGGAGAGAGAGCC 1939  
1858 CCAGCAAGATATGTACAGAACCAATCCAGCTTCAATGCGCAGCAGGTGCGAGAAATCTCTC 1917  
1940 CCAGCAAGATATGTACAGAACCAATCCAGCTTCAATGCGCAGCAGGTGCGAGAAATCTCTC 1999  
1918 CTCAGAGAGAACGCGCATTTATATGTGTGTGAGATGCAAGAAATATGCGCAAGATGTA 1977  
2000 CTCAGAGAGAACGCGCATTTATATGTGTGTGAGATGCAAGAAATATGCGCAAGATGTA 2059  
1978 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGTTTGAAGTTGAAAATCTAAGCAATG 2037  
2060 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGTTTGAAGTTGAAAATCTAAGCAATG 2119  
2038 AAAACCCGTGGCACTTTTAAAGAAAGAAAGCTTACCTTCAAGATTTTGTGATTA 2094  
2120 AAAACCCGTGGCACTTTTAAAGAAAGAAAGCTTACCTTCAAGATTTTGTGATTA 2176

RESULT 3  
US-08-905-223-71  
; Sequence 71, Application US/08905223  
; Patent No. 6222029  
; GENERAL INFORMATION:  
; APPLICANT: Edwards, Jean-Baptiste D.  
; APPLICANT: Duclercq, Aymeric

APPLICANT: Lacroix, Bruno  
TITLE OF INVENTION: 5' ESTs FOR SECRETED PROTEINS  
NUMBER OF SEQUENCES: 503  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe, Martens, Olson & Bear  
STREET: 501 West Broadway  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92101-3505  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: Win95  
SOFTWARE: Word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/905,223  
FILING DATE:  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Israel, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 71:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 390 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: DOUBLE  
TOPOLOGY: LINEAR  
MOLECULE TYPE: CDNA  
ORIGINAL SOURCE:  
ORGANISM: Homo Sapiens  
TISSUE TYPE: Brain  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 289..357  
IDENTIFICATION METHOD: Von Heijne matrix  
OTHER INFORMATION: score 6.9  
OTHER INFORMATION: seq SLSLASHSVSC/SN  
US-08-905-223-71

Query Match 18.4%; Score 386; DB 3; Length 390;  
Best Local Similarity 100.0%; Pred. No. 2.5e-188;  
Matches 386; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 970 GTACAAAGCCTACCTCCAAAGACCTGGAAGATTAAGAGACGCTGCTCTTTG 1029  
DB 3 GTACAAAGCCTACCTCCAAAGACCTGGAAGATTAAGAGACGCTGCTCTTTG 62  
QY 1030 AATAAAGGAGAGACAAAGAGAGAGCTACCTTACCCGACATATACCTGCGGA 1089  
DB 63 AATAAAGGAGAGACAAAGAGAGAGCTACCTTACCCGACATATACCTGCGGA 122  
QY 1090 TGTCTCTCAGTTCATTTTACCTGCTGTGTTGAATCCGAGCAATTCCTAAAGGCA 1149  
DB 123 TGTCTCTCAGTTCATTTTACCTGCTGTGTTGAATCCGAGCAATTCCTAAAGGCA 182  
QY 1150 TTTTGGAGAGCCCTTGTGAGCTATACAGAGACGCTGTAAGAGCCGAGCTACAGAG 1209  
DB 183 TTTTGGAGAGCCCTTGTGAGCTATACAGAGACGCTGTAAGAGCCGAGCTACAGAG 242  
QY 1210 CTGTGAGTAAAGAGGAGGAGCCGATTAATAGCCGCTTTGAGAGATGCTGCTGCTGC 1269  
DB 243 CTGTGAGTAAAGAGGAGGAGCCGATTAATAGCCGCTTTGAGAGATGCTGCTGCTGC 302  
QY 1270 TTGTGATCTGAA 1329  
DB 303 TTGTGATCTGAA 362  
QY 1330 CATCTTCTAACTTCAACCCAGACC 1355

DB 363 CATCTTCTAACTTCAACCCAGACC 388  
RESULT 4  
US-09-949-016-150019  
Sequence 150019, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 150019  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-150019

Query Match 15.8%; Score 330; DB 4; Length 601;  
Best Local Similarity 99.7%; Pred. No. 1.9e-159;  
Matches 380; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 401 GTTAGAATCTGTGTGTTGAGCCGTGATGCTGACTGTGCGCAGCCCTCAGAAACATT 460  
DB 178 GTTAGAATCTGTGTGTTGAGCCGTGATGCTGACTGTGCGCAGCCCTCAGAAACATT 237  
QY 461 TTAGTCAAGCAGAGAGCAAGAGAGATTAAGTGGCCCACTCCCGTGGCATCACTGCAT 520  
DB 238 TTAGTCAAGCAGAGAGCAAGAGAGATTAAGTGGCCCACTCCCGTGGCATCACTGCAT 297  
QY 521 CCTTGGAGCAGACCTTGTGAAGTCAAGCTGCTACACTGAATCAAGTCAAGTCAAGTCA 580  
DB 298 CCTTGGAGCAGACCTTGTGAAGTCAAGCTGCTACACTGAATCAAGTCAAGTCAAGTCA 357  
QY 581 TGAATTCATGATTCAGAGAGAAAGAGATTCAGAGTTTGAAGCAAAATGACAGTAA 640  
DB 358 TGAATTCATGATTCAGAGAGAAAGAGATTCAGAGTTTGAAGCAAAATGACAGTAA 417  
QY 641 GCAACCAATCCAAATGTTGAATTGAAGACTTTGAGTCTCACTTACCCGTTGCGTACCC 700  
DB 418 GCAACCAATCCAAATGTTGAATTGAAGACTTTGAGTCTCACTTACCCGTTGCGTACCC 477  
QY 701 CACTTCAGAGCCTCTGAATATTCCTGTTTACCCCAAGATTTTACAGATCATC 760  
DB 478 CACTTCAGAGCCTCTGAATATTCCTGTTTACCCCAAGATTTTACAGATCATC 537  
QY 761 TGCAGAGTCTCTTGGCCAGG 781  
DB 538 TGCAGAGTCTCTTGGCCAGG 558

RESULT 5  
US-09-949-016-15957  
Sequence 15957, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 15957  
LENGTH: 35916  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-15957

Query Match 15.8%; Score 330; DB 4; Length 35916;  
Best Local Similarity 99.7%; Pred. No. 2.4e-159;  
Matches 380; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 401 GTTTGAACCTTGTTGTTGAGCGGTGATGCTGGAAGCTTGGCCAGCCCTCAGAAAGCATT 460  
DB 10781 GTTTGAACCTTGTTGTTGAGCGGTGATGCTGGAAGCTTGGCCAGCCCTCAGAAAGCATT 10840  
QY 461 TTAGGTCAAGCAGAGCAGAGCAGAGATTAAGTGGCCACTCCCGGTGGATCAGCTGCAT 520  
DB 10841 TTAGGTCAAGCAGAGCAGAGCAGAGATTAAGTGGCCACTCCCGGTGGATCAGCTGCAT 10900  
QY 521 CCTTGAAGCAGAGCCTTGTGAAGTCAAGCTGCTACATGTAATCTCAAGTCAGCTTC 580  
DB 10901 CCTGAGAGCAGAGCCTTGTGAAGTCAAGCTGCTACATGTAATCTCAAGTCAGCTTC 10960  
QY 581 TGAATTCATGATTTAGAGAAAGAAAGATTTGAAGCTTTTGAAGAAATGACGTAAACA 640  
DB 10961 TGAATTCATGATTTAGAGAAAGAAAGATTTGAAGCTTTTGAAGAAATGACGTAAACA 11020  
QY 641 GCAACCAATCCAAATGTTGATTAAGAGCTTGAAGCTTCACTTACCCGTCGGTACCCC 700  
DB 11021 GCAACCAATCCAAATGTTGATTAAGAGCTTGAAGCTTCACTTACCCGTCGGTACCCC 11080  
QY 701 CACTCTCAAGCCTCTCTGATATATTCCTGTTTACCCCGCAATATTTACAGATACATC 760  
DB 11081 CACTCTCAAGCCTCTCTGATATATTCCTGTTTACCCCGCAATATTTACAGATACATC 11140  
QY 761 TGCAGAGTCTCTTGGCCAGG 781  
DB 11141 TGCAGAGTCTCTTGGCCAGG 11161

## RESULT 6

US-09-949-016-150020  
Sequence 150020, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 150020  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-150020

Query Match 13.3%; Score 279; DB 4; Length 601;

Best Local Similarity 99.5%; Pred. No. 3.7e-133;  
Matches 379; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 401 GTTTGAACCTTGTTGTTGAGCGGTGATGCTGGAAGCTTGGCCAGCCCTCAGAAAGCATT 460  
DB 165 GTTTGAACCTTGTTGTTGAGCGGTGATGCTGGAAGCTTGGCCAGCCCTCAGAAAGCATT 224  
QY 461 TTAGGTCAAGCAGAGCAGAGCAGAGATTAAGTGGCCACTCCCGGTGGATCAGCTGCAT 520  
DB 225 TTAGGTCAAGCAGAGCAGAGCAGAGATTAAGTGGCCACTCCCGGTGGATCAGCTGCAT 284  
QY 521 CCTTGAAGCAGAGCCTTGTGAAGTCAAGCTGCTACATGTAATCTCAAGTCAGCTTC 580  
DB 285 CCTGAGAGCAGAGCCTTGTGAAGTCAAGCTGCTACATGTAATCTCAAGTCAGCTTC 344  
QY 581 TGAATTCATGATTTAGAGAAAGAAAGATTTGAAGCTTTTGAAGAAATGACGTAAACA 640  
DB 345 TGAATTCATGATTTAGAGAAAGAAAGATTTGAAGCTTTTGAAGAAATGACGTAAACA 404  
QY 641 GCAACCAATCCAAATGTTGATTAAGAGCTTGAAGCTTCACTTACCCGTCGGTACCCC 700  
DB 405 GCAACCAATCCAAATGTTGATTAAGAGCTTGAAGCTTCACTTACCCGTCGGTACCCC 464  
QY 701 CACTCTCAAGCCTCTCTGATATATTCCTGTTTACCCCGCAATATTTACAGATACATC 760  
DB 465 CACTCTCAAGCCTCTCTGATATATTCCTGTTTACCCCGCAATATTTACAGATACATC 524  
QY 761 TGCAGAGTCTCTTGGCCAGG 781  
DB 525 TGCAGAGTCTCTTGGCCAGG 545

## RESULT 7

US-09-949-016-150037  
Sequence 150037, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 150037  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-150037

Query Match 9.0%; Score 189; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 8.9e-87;  
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1369 AGCTCAAGTTATTTACCCAGAAAGCTCATTTTGTCTTCAACATTTGTGAATTTCTG 1428  
DB 18 AGCTCAAGTTATTTACCCAGAAAGCTCATTTTGTCTTCAACATTTGTGAATTTCTG 77  
QY 1429 TCTACTGCCAACAAGAGTTCTGGGAGGAGATATATACAGGCTGCTGCTGTG 1488  
DB 78 TCTACTGCCAACAAGAGTTCTGGGAGGAGATATATACAGGCTGCTGCTGTG 137  
QY 1489 GTTGCTTCAAGTTCTTCAAGCAATATATATATATATATATATATATATATATATATAT 1548  
DB 138 GTTGCTTCAAGTTCTTCAAGCAATATATATATATATATATATATATATATATATATAT 197

Qy	1549	GCTCCTAAG	1557
Db	198	GCTCCTAAG	206

## RESULT 8

```

US-09-566-921-88
US-Sequence 88, Application US/09566921
Patent No. 6682888
GENERAL INFORMATION:
APPLICANT: Loring, Jeanne F.
APPLICANT: Tingley, Debora W.
APPLICANT: Edwards, Carla M.
TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE
FILE REFERENCE: PA-0024 US
CURRENT APPLICATION NUMBER: US/09/566,921
CURRENT FILING DATE: 2000-05-05
NUMBER OF SEQ ID NOS: 118
SOFTWARE: PERL Program
SEQ ID NO 88
LENGTH: 2475
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. 6682888 255828.26
NAME/KEY: unsure
LOCATION: 1001, 1011
OTHER INFORMATION: a, t, c, g, or other
US-09-566-921-88

```

Query Match	7.5%	Score 158;	DB 4;	Length 2475;
Best Local Similarity	100.0%	Pred. No. 9.2e-71;		
Matches 158;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

Qy	525	GAGGACAGACCTTGTAAAGTCAGAGCTGTACACATTGAATCTCAAGTCAGAGCTTCGAG	584
Db	16	GAGGACAGACCTTGTGAAGTCAGAGCTGTACACATTGAATCTCAAGTCAGAGCTTCGAG	75
Qy	585	ATTTCGATGATTCAGGAAAGAAAGATTCTGAAGTTTGAAGCAAAATGCAGTGAACAGCA	644
Db	76	ATTTCGATGATTCAGGAAAGAAAGATTCTGAAGTTTGAAGCAAAATGCAGTGAACAGCA	135

Qy 645 CCAATCCAATGTTGTAATTGAAGACTTTGAGTCCCTCAC 682  
Db 136 CCAATCCAATGTTGTAATTGAAGACTTTGAGTCCCTCAC 173

## RESULT 5

```

US-09-949-016-150030
Sequence 150030, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIORITY APPLICATION NUMBER: 60/241,755
PRIORITY FILING DATE: 2000-10-20
PRIORITY APPLICATION NUMBER: 60/237,768
PRIORITY FILING DATE: 2000-10-03
PRIORITY APPLICATION NUMBER: 60/231,498
PRIORITY FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 150030
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-150030

```

Query Match	7.4%	Score 155;	DB 4;	Length 601;
Best Local Similarity	100.0%	Pred. No. 3e-69;		
Matches 155;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	904	AATACACACTTTTCCCTATCAGCGCTGGAGATGTCCTTCAGCGCATGCGCTTAACAGTAT	963
Db	320	AATACACACTTTTCCCTATCAGCGCTGGAGATGCTTCAGCGGATCTGCGCTTAACAGTAT	379
Qy	964	TCGAGGATACAAAGCCCTACTCCAAAGACTGAGCGTGAAGATTAAGAAGAGCATCGCGTC	1022
Db	380	TCGAGGATACAAAGCCCTACTCCAAAGACTGAGCGCTGAAGATTAAGAAGAGCATCGCGTC	439
Qy	1024	CTTTTGAAATTAAGGACGACACAAAGACAAAG	1058
Db	440	CTTTTGAAATTAAGGACGACACAAAGACAAAG	474

RESULT 10

```

US-09-949-016-150031
Sequence 150031, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTNER, J. Craig et al.
TITLE OF INVENTION: POLYMERISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE. METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 150031
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-150031

```

Query Match	6.9%;	Score 145;	DB 4;	Length 601;
Best Local Similarity	100.0%;	Pred. No. 4.2e-64;		
Matches 145;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

Qy	904	AATACAGACTTTTCCATCACCCTGGAGATGCGTCAACGCGATCTGCGCTTAACAGTGAT	963
Db	156	AATCAGACTTTTCCATCAGCCTGGAGATGCGTCAACGCGATCTGCGCTTAACAGTGAT	215
Qy	964	TCTTAGGTAACAAGCCTTCTCCAAAGACTGAGCTTGAAGATTAAGAAGACATGCGTC	1022
Db	216	TCTTAGGTAACAAGCCTTCTCCAAAGACTGAGCTTGAAGATTAAGAAGACATGCGTC	275
Qy	1024	CTTTGAAAATAAAGCAGACACAA	1048
Db	276	CTTTGAAAATAAAGCAGACACAA	300

## RESULT 11

```

US-09-949-016-150046
; Sequence 150046, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

```



PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 150046  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-150046

Query Match  
Best Local Similarity 99.5%; Score 137; DB 4; Length 601;  
Pred. No. 5.6e-60;  
Matches 187; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1762 TTCAGAAAAGAGCTCAGACATTTCTTAAGCATGGAGTCTTAATCATCTAAAGTTTCC 1821  
DB 413 TTCAGAAAAGAGCTCAGACATTTCTTAAGCATGGAGTCTTAATCATCTAAAGTTTCC 472  
QY 1822 TTCTCAAGAGATGCTCTGTTGGGAGGAGGAAAGCCCGCAAGATGTAAGCAAC 1881  
DB 473 TTCTCAAGAGATGCTCTGTTGGGAGGAGGAAAGCCCGCAAGATGTAAGCAAC 532  
QY 1882 ATCCAGCTTCATGGCAGAGGATGCGAGATCTCTCCAGAGAGAGCGCATATTAT 1941  
DB 533 ATCCAGCTTCATGGCAGAGGATGCGAGATCTCTCCAGAGAGAGCGCATATTAT 592  
QY 1942 GTGTGTGG 1949  
DB 593 GTGTGTGG 600

## RESULT 12

US-09-949-016-150047  
Sequence 150047, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
PRIOR FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 150047  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-150047

Query Match  
Best Local Similarity 99.5%; Score 137; DB 4; Length 601;  
Pred. No. 5.6e-60;  
Matches 187; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1762 TTCAGAAAAGAGCTCAGACATTTCTTAAGCATGGAGTCTTAATCATCTAAAGTTTCC 1821  
DB 191 TTCAGAAAAGAGCTCAGACATTTCTTAAGCATGGAGTCTTAATCATCTAAAGTTTCC 250  
QY 1822 TTCTCAAGAGATGCTCTGTTGGGAGGAGGAAAGCCCGCAAGATGTAAGCAAC 1881  
DB 251 TTCTCAAGAGATGCTCTGTTGGGAGGAGGAAAGCCCGCAAGATGTAAGCAAC 310  
QY 1882 ATCCAGCTTCATGGCAGAGGATGCGAGATCTCTCCAGAGAGAGCGCATATTAT 1941  
DB 311 ATCCAGCTTCATGGCAGAGGATGCGAGATCTCTCCAGAGAGAGCGCATATTAT 370

QY 1942 GTGTGTGG 1949  
DB 371 GTGTGTGG 378

## RESULT 13

US-09-949-016-150029  
Sequence 150029, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
PRIOR FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 150029  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-150029

Query Match  
Best Local Similarity 100.0%; Score 125; DB 4; Length 601;  
Pred. No. 8.6e-54;  
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 779 AGAGAAAAGCCAGATGATCTGATCTTGCAGAGATCCAGTTTTCAGTCCCAATTCCA 838  
DB 379 AGAGAAAAGCCAGATGATCTGATCTTGCAGAGATCCAGTTTTCAGTCCCAATTCCA 438  
QY 839 AGCGAGTTCACTTCTACGATGATGATGATGATGATGATGATGATGATGATGATGAT 898  
DB 439 AGCGAGTTCACTTCTACGATGATGATGATGATGATGATGATGATGATGATGATGAT 498  
QY 899 TTTCA 903  
DB 499 TTTCA 503

## RESULT 14

US-09-949-016-150041  
Sequence 150041, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
PRIOR FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 150041  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-150041

Query Match  
Best Local Similarity 99.5%; Score 121; DB 4; Length 601;

Best Local Similarity 100.0%; Pred. No. 9.9e-52;  
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1556 AGATATCCATCTCTCTCGAACAACAATTTCTTCCACTTACGATGACCCCTCAATCC 1615  
DB 124 AGATATCCATCTCTCTCGAACAACAATTTCTTCCACTTACGATGACCCCTCAATCC 183

QY 1616 CCATCATTAATGGTGGGTCCAGGAACCGGCATAGCCCGTTATTGGGTTCTTCAACATA 1675  
DB 184 CCATCATTAATGGTGGGTCCAGGAACCGGCATAGCCCGTTATTGGGTTCTTCAACATA 243

QY 1676 G 1676  
DB 244 G 244

## RESULT 15

US-09-949-016-150042  
Sequence 150042, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CU001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 150042  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-150042

Query Match 5.8%; Score 121; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 9.9e-52;

Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1556 AGATATCCATCTCTCTCGAACAACAATTTCTTCCACTTACGATGACCCCTCAATCC 1615

DB 95 AGATATCCATCTCTCTCGAACAACAATTTCTTCCACTTACGATGACCCCTCAATCC 154

QY 1616 CCATCATTAATGGTGGGTCCAGGAACCGGCATAGCCCGTTATTGGGTTCTTCAACATA 1675

DB 155 CCATCATTAATGGTGGGTCCAGGAACCGGCATAGCCCGTTATTGGGTTCTTCAACATA 214

QY 1676 G 1676  
DB 215 G 215

Search completed: August 27, 2005, 16:18:20  
Job time : 237.42 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2005, 00:17:56 ; Search time 899.113 Seconds  
(without alignments)  
15238.608 Million cell updates/sec

Title: US-09-371-347a-45  
Perfect score: 2094  
Sequence: 1 atgaggaggttctgtact.....ttcaggatattggtcattaa 2094

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 7331713 seqs, 3271544945 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 14663426

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Published Applications NA.\*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
20: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
21: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
22: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
23: /cgn2\_6/ptodata/2/pubpna/US11\_PUBCOMB.seq.\*  
24: /cgn2\_6/ptodata/2/pubpna/US11\_PUBCOMB.seq.\*  
25: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
26: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Score	Query Match	Length	ID	Description
No.					
1	2094	100.0	2094	10	US-09-371-347-45
2	1854	88.5	2097	10	US-09-371-347-1
3	1854	88.5	3259	10	US-09-371-347-24
4	1803	86.1	2097	10	US-09-371-347-41
5	1803	86.1	2097	10	US-09-371-347-43
6	1677	80.1	2093	10	US-09-371-347-47
7	896	42.8	3256	21	US-10-741-600-692

8	896	42.8	3274	21	US-10-741-600-693	Sequence 693, App
9	330	15.8	591	16	US-10-029-386-6369	Sequence 6369, App
10	328	15.7	379	16	US-10-029-386-20100	Sequence 20100, A
11	279	13.3	591	16	US-10-029-386-1735	Sequence 1735, App
12	277	13.2	379	16	US-10-029-386-15435	Sequence 15435, A
13	266	12.7	43985	21	US-10-741-600-17757	Sequence 17757, A
14	188	9.0	525	16	US-10-029-386-6533	Sequence 6533, App
15	175	8.4	175	16	US-10-029-386-14338	Sequence 14338, A
16	158	7.5	2475	22	US-09-909-5678-38	Sequence 38, App1
17	158	7.5	2475	22	US-10-765-700-88	Sequence 88, App1
18	158	7.5	21852	21	US-10-741-600-17986	Sequence 17986, A
19	150	7.2	201	21	US-10-741-600-15583	Sequence 15583, A
20	150	7.2	201	21	US-10-741-600-15584	Sequence 15589, A
21	150	7.2	201	21	US-10-741-600-15589	Sequence 15589, A
22	150	7.2	201	21	US-10-741-600-15590	Sequence 15592, A
23	150	7.2	201	21	US-10-741-600-15592	Sequence 15592, A
24	150	7.2	201	21	US-10-741-600-15594	Sequence 15594, A
25	150	7.2	201	21	US-10-741-600-15598	Sequence 15598, A
26	150	7.2	201	21	US-10-741-600-15599	Sequence 15599, A
27	150	7.2	201	21	US-10-741-600-15600	Sequence 15600, A
28	150	7.2	201	21	US-10-741-600-15606	Sequence 15609, A
29	150	7.2	201	21	US-10-741-600-15609	Sequence 15609, A
30	150	7.2	201	21	US-10-741-600-15610	Sequence 15610, A
31	150	7.2	201	21	US-10-741-600-15612	Sequence 15612, A
32	150	7.2	201	21	US-10-741-600-15613	Sequence 15613, A
33	150	7.2	201	21	US-10-741-600-15614	Sequence 15614, A
34	150	7.2	201	21	US-10-741-600-15620	Sequence 15620, A
35	150	7.2	201	21	US-10-741-600-15621	Sequence 15621, A
36	150	7.2	201	21	US-10-741-600-15623	Sequence 15623, A
37	150	7.2	201	21	US-10-741-600-15625	Sequence 15625, A
38	150	7.2	201	21	US-10-741-600-15629	Sequence 15629, A
39	150	7.2	201	21	US-10-741-600-15630	Sequence 15630, A
40	150	7.2	201	21	US-10-741-600-15631	Sequence 15631, A
41	150	7.2	201	21	US-10-741-600-15637	Sequence 15637, A
42	150	7.2	201	21	US-10-741-600-15640	Sequence 15640, A
43	150	7.2	201	21	US-10-741-600-15641	Sequence 15641, A
44	150	7.2	201	21	US-10-741-600-15643	Sequence 15643, A
45	150	7.2	201	21	US-10-741-600-53894	Sequence 53894, A

#### ALIGNMENTS

RESULT 1  
US-09-371-347-45  
; Sequence 45, Application US/09371347  
; Publication No. US20030082676A1  
; GENERAL INFORMATION:  
; APPLICANT: Roy A. Gravel et al.  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE REDUCTASE;  
; TITLE OF INVENTION: CLONING, AND METHODS FOR EVALUATING RISK OF NEURAL TUBE  
; FILE REFERENCE: 50004/003003  
; CURRENT APPLICATION NUMBER: US/09/371,347  
; CURRENT FILING DATE: 1999-08-10  
; PRIOR APPLICATION NUMBER: 60/071,622  
; PRIOR FILING DATE: 1998-01-16  
; PRIOR APPLICATION NUMBER: 09/232,028  
; PRIOR FILING DATE: 1999-01-15  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 45  
; LENGTH: 2094  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-371-347-45

Query Match 100.0%; Score 2094; DB 10; Length 2094;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2094; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 ATGAGGAGTTCTGTACTATATGCTACAGCAGGAGCAGGCAAGGCATCGCAGAA 60  
|||||

D	b	1	ATGAGAGGGTTTCTGTACTATATGTCTACACAGACAGG99ACAGGCAAGGCCATGTGCAGAA	60
Q	y	61	GAATGTGTGAGCAAGCTGTGTGTACATGAGATTTTCTGCAGATCTTCACTGTATTATGAA	120
D	b	61	GAATGTGTGAGCAAGCTGTGTGTACATGAGATTTTCTGCAGATCTTCACTGTATTATGAA	120
Q	y	121	TCCGATTAAGTATGACCTAAACACGAAACAGCTCTCTGTGTGTGTGGATTTCATCCAGC	180
D	b	121	TCCGATTAAGTATGACCTAAACACGAAACAGCTCTCTGTGTGTGTGGATTTCATCCAGC	180
Q	y	181	GGCACCCGAGACCCACCCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACA	240
D	b	181	GGCACCCGAGACCCACCCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACA	240
Q	y	241	CTGCGCGGTGATTTCTTTTGTCTCACTCGCGGTATGGGTATCTGGGTCTCGGTATTCGAA	300
D	b	241	CTGCGCGGTGATTTCTTTTGTCTCACTCGCGGTATGGGTATCTGGGTCTCGGTATTCGAA	300
Q	y	301	TACACTACTTTTGCATATGGGGGGAGATATTAATTAACGACTTCAAGAGCTTGAGCC	360
D	b	301	TACACTACTTTTGCATATGGGGGGAGATATTAATTAACGACTTCAAGAGCTTGAGCC	360
Q	y	361	CGGCAATTTCTATGACACTGACCATGACAGATGACTGTGTAGGTTTGAACCTTGTGTGAG	420
D	b	361	CGGCAATTTCTATGACACTGACCATGACAGATGACTGTGTAGGTTTGAACCTTGTGTGAG	420
Q	y	421	CCGTGATTTGCTGACCTCTGCGCACGCCCTCAAGAAAGCATTTTATGCTCAGCAGAGCAA	480
D	b	421	CCGTGATTTGCTGACCTCTGCGCACGCCCTCAAGAAAGCATTTTATGCTCAGCAGAGCAA	480
Q	y	481	GAGAGATTAAGTGTGCGGACCTCGCGGTGGCATCACTGCATCTCTGAGAGCAGACCTGTG	540
D	b	481	GAGAGATTAAGTGTGCGGACCTCGCGGTGGCATCACTGCATCTCTGAGAGCAGACCTGTG	540
Q	y	541	AAGTCAAGAGCTGTACACATTTGAATCTCAAGTCGAGCTTCTGAGATTCGATTCAGAA	600
D	b	541	AAGTCAAGAGCTGTACACATTTGAATCTCAAGTCGAGCTTCTGAGATTCGATTCAGAA	600
Q	y	601	AGAAAGATTTCTGAGTTTGAAGCAAAATGCACTGAAACAGCAACCAATCAATGTTGTA	660
D	b	601	AGAAAGATTTCTGAGTTTGAAGCAAAATGCACTGAAACAGCAACCAATCAATGTTGTA	660
Q	y	661	ATTGAAGCTTTGAGTCTCTCACTTAACCGTTCGGTACCCCACTCTCAAGCCTCTCTG	720
D	b	661	ATTGAAGCTTTGAGTCTCTCACTTAACCGTTCGGTACCCCACTCTCTCAAGCCTCTCTG	720
Q	y	721	AAATATTCCTGGTTTACCCCGACAGATATTTTACAGGTACATCTGCAGAGATCTTGGCGAG	780
D	b	721	AAATATTCCTGGTTTACCCCGACAGATATTTTACAGGTACATCTGCAGAGATCTTGGCGAG	780
Q	y	781	GAGGAAAGCCAGATCTGTGACTTCAGACAGATCAAGTTTTCAGAGTCCAAATTCGAAG	840
D	b	781	GAGGAAAGCCAGATCTGTGACTTCAGACAGATCAAGTTTTCAGAGTCCAAATTCGAAG	840
Q	y	841	GCAATTCACCTTACTAGATGATGCGCATTAACCACTGTGCTGTGTGAATTTGACATT	900
D	b	841	GCAATTCACCTTACTAGATGATGCGCATTAACCACTGTGCTGTGTGAATTTGACATT	900
Q	y	901	TCAAAATACAGCTTTTCTATCAGCTGAGAGATGCGTTACGCGGATCTGCGCCATACAGT	960
D	b	901	TCAAAATACAGCTTTTCTATCAGCTGAGAGATGCGTTACGCGGATCTGCGCCATACAGT	960
Q	y	961	GATTCGAGGTACAAAGGCTTCTCAAGACTGACGTTGAAGATTAAGAAAGACACTGC	1020
D	b	961	GATTCGAGGTACAAAGGCTTCTCAAGACTGACGTTGAAGATTAAGAAAGACACTGC	1020
Q	y	1021	GTCCTTTTGAATAAAGGACAGACAAAGAAAGAGAGCTTACCTTACCAGCATATA	1080
D	b	1021	GTCCTTTTGAATAAAGGACAGACAAAGAAAGAGAGCTTACCTTACCAGCATATA	1080
Q	y	1081	CTGCGGGAGATCTCTCTCAGCTTCAATTTTACCTGTGTCTTGAATTCGAGCAATTCCT	1140
D	b	1081	CTGCGGGAGATCTCTCTCAGCTTCAATTTTACCTGTGTCTTGAATTCGAGCAATTCCT	1140

Qy	1144	AAAAAGCATTTTGTGCAGCCCTTGTGACATTATACAGTGAACAGTGTGAAAAGCGCAGG	1200
Db	1141	AAAAAGCATTTTGTGCAGCCCTTGTGACATTATACAGTGAACAGTGTGAAAAGCGCAGG	1200
Qy	1201	CTACAGGAGCTGTGACGTAAACAAGGCGACGCCGATTTAAGCCGCTTTGTACAGATGCC	1260
Db	1201	CTACAGGAGCTGTGACGTAAACAAGGCGCGACGCCGATTTAAGCCGCTTTGTACAGATGCC	1260
Qy	1261	TGAGCCCTGTTGTGTGATCTCCCTCCGCTTTCCTTCCTGTCGACGCAACACACAGCTC	1320
Db	1261	TGAGCCCTGTTGTGTGATCTCCCTCCGCTTTCCTTCCTGTCGACGCAACACACAGCTC	1320
Qy	1321	CTGCTCGAACATCTCTCTAACTTCAACCCAGACGATTCGTGTGACAGCTCAAGTTTA	1380
Db	1321	CTGCTCGAACATCTCTCTAACTTCAACCCAGACGATTCGTGTGACAGCTCAAGTTTA	1380
Qy	1381	TTTCAACCAGAAAAGCTCCATTTTGTCTTCAACATTTGTGAAATTTGTCTTACTGCCACA	1440
Db	1381	TTTCAACCAGAAAAGCTCCATTTTGTCTTCAACATTTGTGAAATTTGTCTTACTGCCACA	1440
Qy	1441	ACAGAGGTTGTGGGAAAGGAGATGTACAGGCGGTGGCCCTGTGTTGTTGCTTCAAGTT	1500
Db	1441	ACAGAGGTTGTGGGAAAGGAGATGTACAGGCGGTGGCCCTGTGTTGTTGCTTCAAGTT	1500
Qy	1501	CTTCAGCCAAACATATCATGCAATCCCATGAAAGACAGCGGGAAGACCCCTGCTCCTAAGATA	1560
Db	1501	CTTCAGCCAAACATATCATGCAATCCCATGAAAGACAGCGGGAAGACCCCTGCTCCTAAGATA	1560
Qy	1561	TCCATCTCTCTCGAACAACAAATCTTCTTCACTTACAGATGACCCCTCAATCCCATC	1620
Db	1561	TCCATCTCTCTCGAACAACAAATCTTCTTCACTTACAGATGACCCCTCAATCCCATC	1620
Qy	1621	ATATAGTGTGGTCAAGGAACGGGATGTGCGCCGCTTATTTGGGTTCTTCAACAATPAGAG	1680
Db	1621	ATATAGTGTGGTCAAGGAACGGGATGTGCGCCGCTTATTTGGGTTCTTCAACAATPAGAG	1680
Qy	1681	AAACTCCAAACACACACCAGATGAAATTTTGTGACATGTGTTTGTGCTGCAGG	1740
Db	1681	AAACTCCAAACACACACCAGATGAAATTTTGTGACATGTGTTTGTGCTGCAGG	1740
Qy	1741	CATPAGATATGGGATTTATCTATTCAGAAAAGAGCTCAGACATTTCTTAAAGCATGGGATC	1800
Db	1741	CATPAGATATGGGATTTATCTATTCAGAAAAGAGCTCAGACATTTCTTAAAGCATGGGATC	1800
Qy	1801	TTTAACTCATCTTAAAGGTTCTCTTCTCAAGATGCTCTCTTGTGGAGAGAGAACCCCA	1860
Db	1801	TTTAACTCATCTTAAAGGTTCTCTTCTCAAGATGCTCTCTTGTGGAGAGAGAACCCCA	1860
Qy	1861	GCAAGATATGTACAGACAAATCCAGCTTCATGTGCGACGAGGTGGGAGAAATCTCTC	1920
Db	1861	GCAAGATATGTACAGACAAATCCAGCTTCATGTGCGACGAGGTGGGAGAAATCTCTC	1920
Qy	1921	CAGAGAAACGGCCATATTTATGTGTGTGAGATGCAAAAGATATGTGCAAGATGTACAT	1980
Db	1921	CAGAGAAACGGCCATATTTATGTGTGTGAGATGCAAAAGATATGTGCAAGATGTACAT	1980
Qy	1981	GATGCCCTTGTGCAAAATPATAAGCAAAAGAGTTGAGTTGAAATACTAGAAAGCAATGAAA	2040
Db	1981	GATGCCCTTGTGCAAAATPATAAGCAAAAGAGTTGAGTTGAAATACTAGAAAGCAATGAAA	2040
Qy	2041	ACCCTGGCACTTTAAAGAAAGAAAACGCTACCTTCAGATATTTGTCATTA	2094
Db	2041	ACCCTGGCACTTTAAAGAAAGAAAACGCTACCTTCAGATATTTGTCATTA	2094
RESULT 2			
US-09-371-347-1			
; Sequence 1, Application US/09371347			
; Publication No. US2003002676A1			
; GENERAL INFORMATION: Roy A. Gravel et al.			
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE REDUCTASE:			

TITLE OF INVENTION: CLONING, AND METHODS FOR EVALUATING RISK OF NEURAL TUBE  
DEFECTS, CARDIOVASCULAR DISEASE, AND CANCER  
FILE REFERENCE: 50004/003003  
CURRENT APPLICATION NUMBER: US/09/371,347  
CURRENT FILING DATE: 1999-08-10  
PRIOR APPLICATION NUMBER: 60/071,622  
PRIOR FILING DATE: 1998-01-16  
PRIOR APPLICATION NUMBER: 09/232,028  
PRIOR FILING DATE: 1999-01-15  
NUMBER OF SEQ ID NOS: 51  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 2097  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-371-347-1

Query Match 88.5%; Score 1854; DB 10; Length 2097;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2094; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 ATGAGAGGTTCTGTTACTATATGCTACACAGGAGGAGCAAGGCAAGCCATCGAGAA 60  
DB 1 ATGAGAGGTTCTGTTACTATATGCTACACAGGAGGAGCAAGGCAAGCCATCGAGAA 60  
QY 61 GAAATGTGACCAAGCTGTGTACATGATTTCTGAGACTTCACTGTATTAGTGA 120  
DB 61 GAAATGTGACCAAGCTGTGTACATGATTTCTGAGACTTCACTGTATTAGTGA 120  
QY 121 TCCGATAAGTATGACCTTAATAACCGAAGAGCTCTCTGTTGTTGTTGTTCTACAG 180  
DB 121 TCCGATAAGTATGACCTTAATAACCGAAGAGCTCTCTGTTGTTGTTGTTCTACAG 180  
QY 181 GGCACCGGAGACCCACCCGACACAGCCGCAAGTTGTTAAGAAATACAGAACCAACA 240  
DB 181 GGCACCGGAGACCCACCCGACACAGCCGCAAGTTGTTAAGAAATACAGAACCAACA 240  
QY 241 CTGCGGTTGATTTCTTCTGCTACCTGCGGTATGAGTTTCTGAGTTTCTGAGTAA 300  
DB 241 CTGCGGTTGATTTCTTCTGCTACCTGCGGTATGAGTTTCTGAGTTTCTGAGTAA 300  
QY 301 TACACCTACTTTTGCAATGAGGAGGAGATTAATTGATTAACGACTTCAAGAGCTTGAGCC 360  
DB 301 TACACCTACTTTTGCAATGAGGAGGAGATTAATTGATTAACGACTTCAAGAGCTTGAGCC 360  
QY 361 CGGCAATTTCTATGACACTGAGACATGCAATGACTGTGAGTTTGAACCTTGTTGAG 420  
DB 361 CGGCAATTTCTATGACACTGAGACATGCAATGACTGTGAGTTTGAACCTTGTTGAG 420  
QY 421 CCGTGAATGCTGAGACTCTGCGGACGCTTCAAAAAGCAATTTTGTGAGAGAGAGCA 480  
DB 421 CCGTGAATGCTGAGACTCTGCGGACGCTTCAAAAAGCAATTTTGTGAGAGAGAGCA 480  
QY 481 GAGGAGATAGTGGGCACTCCCGGTGGGATCACTGCACTCTTGAGAGAGAGCTTGTG 540  
DB 481 GAGGAGATAGTGGGCACTCCCGGTGGGATCACTGCACTCTTGAGAGAGAGCTTGTG 540  
QY 541 AAGTCAGAGCTGCTACATGATCTCAAGTCAAGCTTCTGAGATTGATGATTCAGAA 600  
DB 541 AAGTCAGAGCTGCTACATGATCTCAAGTCAAGCTTCTGAGATTGATGATTCAGAA 600  
QY 601 AGAAGAGATTCTGAGGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATTCATGTTGA 660  
DB 601 AGAAGAGATTCTGAGGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATTCATGTTGA 660  
QY 661 ATTGAAGATTGAGCTCTACCTTACCCGTTGGTATCCCACTCTCAAGGCTCTCTG 720  
DB 661 ATTGAAGATTGAGCTCTACCTTACCCGTTGGTATCCCACTCTCTCAAGGCTCTCTG 720  
QY 721 AATATTCCTGTTTACCCCAAGATATTTACAGATCATCTGAGAGAGTCTCTGAGCAG 780  
DB 721 AATATTCCTGTTTACCCCAAGATATTTACAGATCATCTGAGAGAGTCTCTGAGCAG 780

QY 781 GAGGAAAGCAAGTATCTGATCTTACAGAGATCCAGTTTTCAGATGCCAATTTCAAG 840  
DB 781 GAGGAAAGCAAGTATCTGATCTTACAGAGATCCAGTTTTCAGATGCCAATTTCAAG 840  
QY 841 GCAATTCATCTTACAGATATGATCCATTAACCACTCTGCTGTGATGATTTGACAT 900  
DB 841 GCAATTCATCTTACAGATATGATCCATTAACCACTCTGCTGTGATGATTTGACAT 900  
QY 901 TCAATACAGACTTTTCTATGAGCTGAGATGCTTCAAGGATATCTGCTTACAGAT 960  
DB 901 TCAATACAGACTTTTCTATGAGCTGAGATGCTTCAAGGATATCTGCTTACAGAT 960  
QY 961 GATTCGAGTACCAAGCTTCTCAAGAGCTGACCTTGAAGATTAAGAGAGAGCTG 1020  
DB 961 GATTCGAGTACCAAGCTTCTCAAGAGCTGACCTTGAAGATTAAGAGAGAGCTG 1020  
QY 1021 GTCTTTGAAATTAAGGAGACACAAAGAAAGAGAGTACCTTACCCAGCATTA 1080  
DB 1021 GTCTTTGAAATTAAGGAGACACAAAGAAAGAGAGTACCTTACCCAGCATTA 1080  
QY 1081 CCGGAGGAGATGTTCTTCAAGTTTCTTACCTGAGTCTGAAATCCGAGCAATCT 1140  
DB 1081 CCGGAGGAGATGTTCTTCAAGTTTCTTACCTGAGTCTGAAATCCGAGCAATCT 1140  
QY 1141 AAAAAGCATTTTTCGAGCCCTTGTGACTATACAGTGAAGTCTGAAAAGGCGCAG 1200  
DB 1141 AAAAAGCATTTTTCGAGCCCTTGTGACTATACAGTGAAGTCTGAAAAGGCGCAG 1200  
QY 1201 CTACAGAGCTGTGAGTAAACAGAGGAGCCGATTAATAGCCGCTTGTGAGAGTCC 1260  
DB 1201 CTACAGAGCTGTGAGTAAACAGAGGAGCCGATTAATAGCCGCTTGTGAGAGTCC 1260  
QY 1261 TGTGCTGTTTGTGATCTCTCTGCTGCTTCCCTTCTGAGCAGCAGCAGCTGCTC 1320  
DB 1261 TGTGCTGTTTGTGATCTCTCTGCTGCTTCCCTTCTGAGCAGCAGCAGCTGCTC 1320  
QY 1321 CTGCTGAAATCTTCTTAACTTCAACCCAGACATATTCGTGACAGCTCAAGTTTA 1380  
DB 1321 CTGCTGAAATCTTCTTAACTTCAACCCAGACATATTCGTGACAGCTCAAGTTTA 1380  
QY 1381 TTTCAACCAAGAAAGCTCAATTTTGTCAATTTGTAATTTCTGCTACCTGACCA 1440  
DB 1381 TTTCAACCAAGAAAGCTCAATTTTGTCAATTTGTAATTTCTGCTACCTGACCA 1440  
QY 1441 ACAAGGTTCTGCGGAGGAGATGATGATGAGCTGCTGCTGTTGTTGTTGAT 1500  
DB 1441 ACAAGGTTCTGCGGAGGAGATGATGATGAGCTGCTGCTGTTGTTGTTGAT 1500  
QY 1501 CTTGAGCCAAATATCATGATCCCATGAAAGACGAGGAGAAAGCCCTGCTCTAAGATA 1560  
DB 1501 CTTGAGCCAAATATCATGATCCCATGAAAGACGAGGAGAAAGCCCTGCTCTAAGATA 1560  
QY 1561 TCCATCTCTCTGAGACAAATTTCTTCACTTACAGATGACCCCTCAATCCCCTATC 1620  
DB 1561 TCCATCTCTCTGAGACAAATTTCTTCACTTACAGATGACCCCTCAATCCCCTATC 1620  
QY 1621 ATTAATGTTGTTGTCAGAAACCGGATAGCCGCTTATTTGAGTTCTTACAACTAAGAG 1680  
DB 1621 ATTAATGTTGTTGTCAGAAACCGGATAGCCGCTTATTTGAGTTCTTACAACTAAGAG 1680  
QY 1681 AAATCTCAAGAACCAACACAGATGAAATTTTGGAGCAATGNG--GTTTTTGGCTGC 1737  
DB 1681 AAATCTCAAGAACCAACACAGATGAAATTTTGGAGCAATGNG--GTTTTTGGCTGC 1737  
QY 1738 AGGCAATAGATAGGATTAATCTATTCAGAAAAGCTCAGCAATTCCTTAAGATGG 1797  
DB 1738 AGGCAATAGATAGGATTAATCTATTCAGAAAAGCTCAGCAATTCCTTAAGATGG 1797  
QY 1741 AGGCAATAGATAGGATTAATCTATTCAGAAAAGCTCAGCAATTCCTTAAGATGG 1800  
DB 1741 AGGCAATAGATAGGATTAATCTATTCAGAAAAGCTCAGCAATTCCTTAAGATGG 1800  
QY 1798 ATCTTAATCTATTAAGGTTCTTCTCAAGATGCTCTGTTGGAGAGAGAGGCT 1857  
DB 1798 ATCTTAATCTATTAAGGTTCTTCTCAAGATGCTCTGTTGGAGAGAGAGGCT 1857  
QY 1801 ATCTTAATCTATTAAGGTTCTTCTCAAGATGCTCTGTTGGAGAGAGAGGCT 1860  
DB 1801 ATCTTAATCTATTAAGGTTCTTCTCAAGATGCTCTGTTGGAGAGAGAGGCT 1860  
QY 1858 CCAAGCAAGTATGTAACAAGCAATTCAGCTTCAATGTCAGAGAGTGGAGAAATCTTC 1917

1861 CCGAGGAAGATGATGACAAAGCAATCAGCTTCATGCGACAGAGGGGAGAAATCTTC 1920  
1918 CTCACAGAGAACGGCCATATTATGTTGTGTGAGATGCAAAAGATATGCGCAAGATGTA 1977  
1921 CTCACAGAGAACGGCCATATTATGTTGTGTGAGATGCAAAAGATATGCGCAAGATGTA 1980  
1978 CATGATGCCCTTGTGCAATTAATTAAGCAAAAGAGTTGAGTTGAAAACTAGAAGCAATG 2037  
1981 CATGATGCCCTTGTGCAATTAATTAAGCAAAAGAGTTGAGTTGAAAACTAGAAGCAATG 2040  
2038 AAAACCTGGCGCCTTTAAAAAGAAAGCAAAAGCCTTACAGATATTTGTCATTA 2094  
2041 AAAACCTGGCGCCTTTAAAAAGAAAGCAAAAGCCTTACAGATATTTGTCATTA 2097

RESULT 3  
US-09-371-347-24  
Sequence 24, Application US/09371347  
Publication No. US20030082676A1  
GENERAL INFORMATION:  
APPLICANT: Roy A. Gravel et al.  
TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE REDUCTASE:  
TITLE OF INVENTION: CLONING, AND METHODS FOR EVALUATING RISK OF NEURAL TUBE  
TITLE OF INVENTION: DEFECTS, CARDIOVASCULAR DISEASE, AND CANCER  
FILE REFERENCE: 50004/003003  
CURRENT APPLICATION NUMBER: US/09/371.347  
CURRENT FILING DATE: 1999-08-10  
PRIOR APPLICATION NUMBER: 60/071.622  
PRIOR FILING DATE: 1998-01-16  
PRIOR APPLICATION NUMBER: 09/232.028  
PRIOR FILING DATE: 1999-01-15  
NUMBER OF SEQ ID NOS: 51  
SOFTWARE: PaatSeq for windows Version 4.0  
SEQ ID NO 24  
LENGTH: 3259  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-371-347-24

Query Match 88.5%; Score 1854; DB 10; Length 3259;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2094; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

1 ATGAGAGGTTTCTGTTACTATATGCTACACAGCAGGAGCAAGCAAGGCATGCGAGAA 60  
80 ATGAGAGGTTTCTGTTACTATATGCTACACAGCAGGAGCAAGCAAGGCATGCGAGAA 139  
61 GAATGTTGAGCAAGCTGTGTGATCATGATTTTCTGCAATCTTCACTGATTAAGTAA 120  
140 GAATGTTGAGCAAGCTGTGTGATCATGATTTTCTGCAATCTTCACTGATTAAGTAA 199  
121 TCCGATTAAGTATGACTATAAACCAGAAAGCTCTCTTGTGTGTGTGTTTCTACCAAG 180  
200 TCCGATTAAGTATGACTATAAACCAGAAAGCTCTCTTGTGTGTGTGTTTCTACCAAG 259  
181 GGCAACCGGAGACCAACCCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACA 240  
260 GGCAACCGGAGACCAACCCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACA 319  
241 CTGCGCGGTGATTTCTTGTGCTCAACCTGCGGTGATGAGTTTCTGCGTCTGCGGTGATTCAGAA 300  
320 CTGCGCGGTGATTTCTTGTGCTCAACCTGCGGTGATGAGTTTCTGCGTCTGCGGTGATTCAGAA 379  
301 TACACCTACTTTTGGCAATGGGGGGAAGATATTAAGCACTTCAAGAGCTTGGAGCC 360  
380 TACACCTACTTTTGGCAATGGGGGGAAGATATTAAGCACTTCAAGAGCTTGGAGCC 439  
361 CGGCAATTTCTATGACACTGACATGACATGACATGCTGTGTAGGTTTGAACCTTGTGTGAG 420  
440 CGGCAATTTCTATGACACTGACATGACATGACATGCTGTGTAGGTTTGAACCTTGTGTGAG 499  
421 CCGTGATTTGCTGACCTCTGGCCAGCCCTCAGAAAGCAATTTTGTAGGTACAGACAGAGCAAA 480

500 CCGTGATTTGCTGACCTCTGGCCAGCCCTCAGAAAGCAATTTTGTAGGTACAGACAGAGCAAA 559  
481 GAGAGATTAAGTGGCGCACTCCCGGTGGATCATCTGCAATCTTGAAGACAGACCTTGTG 540  
560 GAGAGATTAAGTGGCGCACTCCCGGTGGATCATCTGCAATCTTGAAGACAGACCTTGTG 619  
541 AAGTCAGAGCTGTCACATTTGAATCTCAAGTCAAGCTTCTGAGATTCATGATTCAGAA 600  
620 AAGTCAGAGCTGTCACATTTGAATCTCAAGTCAAGCTTCTGAGATTCATGATTCAGAA 679  
601 AGAAAGATTTCTGAGGTTTGAAGCAAAATGCAAGTAAAGCAACCAATCCAAATGTTGTA 660  
680 AGAAAGATTTCTGAGGTTTGAAGCAAAATGCAAGTAAAGCAACCAATCCAAATGTTGTA 739  
661 ATTGAAGACTTTGAGTCTCATCTTACCCGTTGGATCCCGCACTCTCAAGAGCTTCTG 720  
740 ATTGAAGACTTTGAGTCTCATCTTACCCGTTGGATCCCGCACTCTCAAGAGCTTCTG 799  
721 AATATTCCTGTTTACCCCGAGAAATTTTACAGTACATCTGCAAGAGTCTCTTGGCCAG 859  
800 AATATTCCTGTTTACCCCGAGAAATTTTACAGTACATCTGCAAGAGTCTCTTGGCCAG 859  
781 GAGGAAGCCAGATCTGTGACTTACAGAGATCCAGTTTCAAGTGCATTTGAAG 840  
860 GAGGAAGCCAGATCTGTGACTTACAGAGATCCAGTTTCAAGTGCATTTGAAG 919  
841 GCAGTTCAACTTACTACGATGATGCAATTAACCACTCTGCTGTAGATTTGACAT 900  
920 GCAGTTCAACTTACTACGATGATGCAATTAACCACTCTGCTGTAGATTTGACAT 979  
901 TCAATATCAGACTTTTCTATCAGCTTGAAGATGCTTCAAGCTGATCTGCCATACAGT 960  
980 TCAATATCAGACTTTTCTATCAGCTTGAAGATGCTTCAAGCTGATCTGCCATACAGT 1039  
961 GATTCGAGGTACAAAGCCATCTCCAAAGCTGACCTTGAAGATTAAGAGAGCACTGC 1020  
1040 GATTCGAGGTACAAAGCCATCTCCAAAGCTGACCTTGAAGATTAAGAGAGCACTGC 1099  
1021 GTCTTTTGAATAAATAAGGACAGACAAAGAAAGAGACTTACTTACCCGACATATA 1080  
1100 GTCTTTTGAATAAATAAGGACAGACAAAGAAAGAGACTTACTTACCCGACATATA 1159  
1081 CTTGCGGAGATGTTCTCTCCAGTTCAATTTTACCTGCTGTGAAATCCAGCAATTCCT 1140  
1160 CTTGCGGAGATGTTCTCTCCAGTTCAATTTTACCTGCTGTGAAATCCAGCAATTCCT 1219  
1141 AAAAAGCAATTTTGGAGCCCTTGGACCTATACAGTGAACAGTGTGAAAAGCGCAGG 1200  
1220 AAAAAGCAATTTTGGAGCCCTTGGACCTATACAGTGAACAGTGTGAAAAGCGCAGG 1279  
1201 CTACAGAGCTGTGACATTAACAAAGGGGAGCCGATTAATAGCCCTTGTATCAGATGCGC 1260  
1280 CTACAGAGCTGTGACATTAACAAAGGGGAGCCGATTAATAGCCCTTGTATCAGATGCGC 1339  
1261 TGTGCTGCTGTTGGATCTCTCTGCTGTTTCCCTTCTTGGACGACCACTCAAGTCTC 1320  
1340 TGTGCTGCTGTTGGATCTCTCTGCTGTTTCCCTTCTTGGACGACCACTCAAGTCTC 1399  
1321 CTGCTGGAACATTTCTTAACCTTCAACCGACATTAATTCGTTGCAAGCTCAAGTTTA 1380  
1400 CTGCTGGAACATTTCTTAACCTTCAACCGACATTAATTCGTTGCAAGCTCAAGTTTA 1459  
1381 TTTCAACCCAGAAAGCTCCATTTTGTCTTCAACATTTGGAATTTCTGTCTACGACACA 1440  
1460 TTTCAACCCAGAAAGCTCCATTTTGTCTTCAACATTTGGAATTTCTGTCTACGACACA 1519  
1441 ACAGAGTTCTGCGGAAGGAGATATTAAGCTGCTGCTGCTGTTGTTGTTGCTTCAATT 1500  
1520 ACAGAGTTCTGCGGAAGGAGATATTAAGCTGCTGCTGCTGTTGTTGTTGCTTCAATT 1579  
1501 CTTCAAGCCAAACATTAATCATCTCCATGGAAGCAGGCGGAAAGCCCTGCTCTTAAGATA 1560

Db 1580 CTTGACCAACATACATGATCCCATGAGAGAGAGGAGAGCCCTGCTCTTAAGATA 1639  
Qy 1561 TCCATCTCTCTGGAACAATTTCTTCACTTACGATGACCCCTCAATCCCATC 1620  
Db 1640 TCCATCTCTCTGGAACAATTTCTTCACTTACGATGACCCCTCAATCCCATC 1699  
Qy 1621 ATATGTTGGTTCAGAGAACCCGATAGCCCGTTTATTTGGTTCCTCAACATAGAGAG 1680  
Db 1700 ATATGTTGGTTCAGAGAACCCGATAGCCCGTTTATTTGGTTCCTCAACATAGAGAG 1759  
Qy 1681 AAATCTCAAGAACACACCCAGATGGAATTTTGGAGCATGTG--GTTTTTGGCTGC 1737  
Db 1760 AAATCTCAAGAACACACCCAGATGGAATTTTGGAGCATGTGTGTTTTTGGCTGC 1819  
Qy 1738 AGGCAATAGAGATAGGATATCTATTCAGAAAAGCTCAGACATTTCTTAAAGCATGAG 1797  
Db 1820 AGGCAATAGAGATAGGATATCTATTCAGAAAAGCTCAGACATTTCTTAAAGCATGAG 1879  
Qy 1798 ATCTTAATCTCAATTAAGGTTTCTTCTCAAGAGATGCTCTGTTGGAGAGAGAGCC 1857  
Db 1880 ATCTTAATCTCAATTAAGGTTTCTTCTCAAGAGATGCTCTGTTGGAGAGAGAGCC 1939  
Qy 1858 CCAGCAAGATATGTAACAAGACATCCAGCTTCATGCGCAGAGAGTGGCAGAAATCTTC 1917  
Db 1940 CCAGCAAGATATGTAACAAGACATCCAGCTTCATGCGCAGAGAGTGGCAGAAATCTTC 1999  
Qy 1918 CTCACGAGAGAGGCAATTTTATGTTGTGAGATGCAAAATATGCGCAAGATGTA 1977  
Db 2000 CTCACGAGAGAGGCAATTTTATGTTGTGAGATGCAAAATATGCGCAAGATGTA 2059  
Qy 1978 CATGATGCTCTGTGCAATTAATTAAGCAAGAGGTTGAGTTGAAAACCTAGAGCATG 2037  
Db 2060 CATGATGCTCTGTGCAATTAATTAAGCAAGAGGTTGAGTTGAAAACCTAGAGCATG 2119  
Qy 2038 AAAACCTTGCCCATTTTAAAGAGAAAAACCTACCTTCAGAGATTTTGTCTATA 2094  
Db 2120 AAAACCTTGCCCATTTTAAAGAGAAAAACCTACCTTCAGAGATTTTGTCTATA 2176

RESULT 4  
US-09-371-347-41  
; Sequence 41, Application US/09371347  
; Publication No. US20030082676A1  
; GENERAL INFORMATION:  
; APPLICANT: Roy A. Gravel et al.  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE REDUCTASE.  
; TITLE OF INVENTION: CLONING, AND METHODS FOR EVALUATING RISK OF NEURAL TUBE  
; FILE REFERENCE: 50004/003003  
; CURRENT APPLICATION NUMBER: US/09/371,347  
; PRIOR FILING DATE: 1999-08-10  
; PRIOR APPLICATION NUMBER: 60/071,622  
; PRIOR FILING DATE: 1998-01-16  
; PRIOR APPLICATION NUMBER: 09/232,028  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 41  
; LENGTH: 2097  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-371-347-41

Query Match 86.1%; Score 1803; DB 10; Length 2097;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2093; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

Qy 1 ATGAGAGAGTTCTGTTACTATATGCTACAGCAGAGGACAGGCAAAAGGCATCGAGAA 60  
Db 1 ATGAGAGAGTTCTGTTACTATATGCTACAGCAGAGGACAGGCAAAAGGCATCGAGAA 60  
Qy 61 GAAATGTGAGCAAGCTGTGATCATGATTTTCTGAGATCTCATGTATTAGTGA 120  
Db 61 GAAATGTGAGCAAGCTGTGATCATGATTTTCTGAGATCTCATGTATTAGTGA 120

Db 61 GAAATGTGAGCAAGCTGTGATCATGATTTTCTGAGATCTCATGTATTAGTGA 120  
Qy 121 TCCGATTAAGTATGACCTTAAAAACCAAGCAAGCTCTCTGTTGTGTTCTACACAG 180  
Db 121 TCCGATTAAGTATGACCTTAAAAACCAAGCAAGCTCTCTGTTGTGTTCTACACAG 180  
Qy 181 GGCACCGGAGACCCACCCGACACAGCCCGCAAGTTTGTTAAGAAATACAGAACCAACA 240  
Db 181 GGCACCGGAGACCCACCCGACACAGCCCGCAAGTTTGTTAAGAAATACAGAACCAACA 240  
Qy 241 CTGCGGTGATTTCTTTGCTCACTGCGGTATGAGTTATCGGGTCTCGGTATTCAGAA 300  
Db 241 CTGCGGTGATTTCTTTGCTCACTGCGGTATGAGTTATCGGGTCTCGGTATTCAGAA 300  
Qy 301 TACACCTACTTTTGAATGAGGAGAGATTAATGATTAACGACTTCAAGAGCTTGAGGC 360  
Db 301 TACACCTACTTTTGAATGAGGAGAGATTAATGATTAACGACTTCAAGAGCTTGAGGC 360  
Qy 361 CGGCAATTTCTATGACACTGACATGACATGACATGACATGACATGACATGACATGAC 420  
Db 361 CGGCAATTTCTATGACACTGACATGACATGACATGACATGACATGACATGACATGAC 420  
Qy 421 CCGTGATGCTGACCTGCGCAGACCTCAGAAAGCATTTTAAAGTCAAGAGAGACA 480  
Db 421 CCGTGATGCTGACCTGCGCAGACCTCAGAAAGCATTTTAAAGTCAAGAGAGACA 480  
Qy 481 GAGGAGATAGTGGGCACTCCCGGTGATACCTGATGAGTCTGATGAGAGAGAGAGAG 540  
Db 481 GAGGAGATAGTGGGCACTCCCGGTGATACCTGATGAGTCTGATGAGAGAGAGAGAG 540  
Qy 541 AAGTCAAGCTCTACACATTAATCTCAATGAGCTTCTGAGATTCATGATTCAGAG 600  
Db 541 AAGTCAAGCTCTACACATTAATCTCAATGAGCTTCTGAGATTCATGATTCAGAG 600  
Qy 601 AGAAAGATTCGAGGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATTCATGTTGA 660  
Db 601 AGAAAGATTCGAGGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATTCATGTTGA 660  
Qy 661 ATTGAAGATTTGAGTCTCACTTACCTGCTGATGAGTCTGATGAGTCTGATGAGTCTG 720  
Db 661 ATTGAAGATTTGAGTCTCACTTACCTGCTGATGAGTCTGATGAGTCTGATGAGTCTG 720  
Qy 721 AATATTCCTGTTTACCCCAAGATATTTTACAGTACATCTGAGAGTCTCTGCGCAG 780  
Db 721 AATATTCCTGTTTACCCCAAGATATTTTACAGTACATCTGAGAGTCTCTGCGCAG 780  
Qy 781 GAGGAAAGCAAGTATCTGTGACTTCAGAGATTCAGAGTCTTCAAGTCCAAATTCAGAG 840  
Db 781 GAGGAAAGCAAGTATCTGTGACTTCAGAGATTCAGAGTCTTCAAGTCCAAATTCAGAG 840  
Qy 841 GCAATTCATTAATCAAGATGATGCAATTAACCACTCTGCTGTATTAATTTGACAT 900  
Db 841 GCAATTCATTAATCAAGATGATGCAATTAACCACTCTGCTGTATTAATTTGACAT 900  
Qy 901 TCAAAATACAGACTTTTCTATCAGCTGAGAGTCTTCAAGGATCTGAGGATCTGAGG 960  
Db 901 TCAAAATACAGACTTTTCTATCAGCTGAGAGTCTTCAAGGATCTGAGGATCTGAGG 960  
Qy 961 GATTCGAGTACAAAGCTCTCAAGACTGACCTTGAAGATTAAGAGAGAGAGAGAGAG 1020  
Db 961 GATTCGAGTACAAAGCTCTCAAGACTGACCTTGAAGATTAAGAGAGAGAGAGAGAG 1020  
Qy 1021 GTTCCTTTGAAATTAAGGACACACAAAGAAAGAGAGTACCTTACCCGACATTA 1080  
Db 1021 GTTCCTTTGAAATTAAGGACACACAAAGAAAGAGAGTACCTTACCCGACATTA 1080  
Qy 1081 CCGCGGAGATTTCTCTCAGTTCAATTTTACCTGCTGCTTGAATCCGAGCAATTCCT 1140  
Db 1081 CCGCGGAGATTTCTCTCAGTTCAATTTTACCTGCTGCTTGAATCCGAGCAATTCCT 1140  
Qy 1141 AAAAGGCAATTTTGGAGCCCTTGTGACATACAGTGAACAGTGTGAAAAGGCGAG 1200  
Db 1141 AAAAGGCAATTTTGGAGCCCTTGTGACATACAGTGAACAGTGTGAAAAGGCGAG 1200



QY 1201 CTACAGAGCTGTGCACTAAACAAGGGGAGCCGATTAATAGCCGTTTGTACAGATGCC 1260  
DB 1201 CTACAGAGCTGTGCACTAAACAAGGGGAGCCGATTAATAGCCGTTTGTACAGATGCC 1260  
QY 1261 TGTGCGCTGTGTGATGATCTCCCTCTCGCTTTCCCTTCCGACGACCACTCACTTC 1320  
DB 1261 TGTGCGCTGTGTGATGATCTCCCTCTCGCTTTCCCTTCCGACGACCACTCACTTC 1320  
QY 1321 CTGCTGGAACATCTTCTTAACCTTCAACCCAGACCAATATTCGTGCAAGCTCAAGTTTA 1380  
DB 1321 CTGCTGGAACATCTTCTTAACCTTCAACCCAGACCAATATTCGTGCAAGCTCAAGTTTA 1380  
QY 1381 TTTACCCAGAGAAAGCTCCATTTTGTCTTCAACATTTGTGGAATTTTGTCTACTGCCACA 1440  
DB 1381 TTTACCCAGAGAAAGCTCCATTTTGTCTTCAACATTTGTGGAATTTTGTCTACTGCCACA 1440  
QY 1441 ACAGAGTTTGTGGGAGAGGAGATGATGACAGCTGTGCTGTGTTGTGTTGTTCAATT 1500  
DB 1441 ACAGAGTTTGTGGGAGAGGAGATGATGACAGCTGTGCTGTGTTGTGTTGTTCAATT 1500  
QY 1501 CTTCAGCCAAACATACATGATCCCATGAGAGAGCGGAAAGCCCTGCTCCTTAAGATA 1560  
DB 1501 CTTCAGCCAAACATACATGATCCCATGAGAGAGCGGAAAGCCCTGCTCCTTAAGATA 1560  
QY 1561 TCCATCTCTCTGGAACAACAATTTCTTTCACCTTACAGATGACCCCTCAATCCCCTATC 1620  
DB 1561 TCCATCTCTCTGGAACAACAATTTCTTTCACCTTACAGATGACCCCTCAATCCCCTATC 1620  
QY 1621 ATATAGTGTGCTCAAGAAACCGGCAATGCCCCGTTTATTTGGGTTCTTCAACAATAGAG 1680  
DB 1621 ATATAGTGTGCTCAAGAAACCGGCAATGCCCCGTTTATTTGGGTTCTTCAACAATAGAG 1680  
QY 1681 AAATCTCAAGAAACAACCCAGATGGAATTTTGGAGCAATGTG---GTTTTTGGCTGC 1737  
DB 1681 AAATCTCAAGAAACAACCCAGATGGAATTTTGGAGCAATGTG---GTTTTTGGCTGC 1737  
QY 1738 AGGCAATAGAGTATGAGATTAATCTATTCAGAAAAGAGCTCAGACATTTCTTAAAGATGG 1797  
DB 1738 AGGCAATAGAGTATGAGATTAATCTATTCAGAAAAGAGCTCAGACATTTCTTAAAGATGG 1797  
QY 1798 ATCTTAATCATTAAGATTTCTCTCAAGAGATGCTCTGTGGGGAGAGAGAGCC 1857  
DB 1798 ATCTTAATCATTAAGATTTCTCTCAAGAGATGCTCTGTGGGGAGAGAGAGCC 1857  
QY 1801 ATCTTAATCATTAAGATTTCTCTCAAGAGATGCTCTGTGGGGAGAGAGAGCC 1860  
DB 1801 ATCTTAATCATTAAGATTTCTCTCAAGAGATGCTCTGTGGGGAGAGAGAGCC 1860  
QY 1858 CCAGCAAGATATACAGACAATCAGCTTCATGSCCAGAGAGTGGAGAAATCTTC 1917  
DB 1858 CCAGCAAGATATATGACAGACAATCAGCTTCATGSCCAGAGAGTGGAGAAATCTTC 1917  
QY 1918 CTCAGAGAGAGGAGCTATTTATGTGTGTGAGATGCAAAAGATATGSCCAAGATGTA 1977  
DB 1918 CTCAGAGAGAGGAGCTATTTATGTGTGTGAGATGCAAAAGATATGSCCAAGATGTA 1977  
QY 1978 CATGATGCTCTGTGCAAAATTAATGAGAGAGTGGAGTGGAAAACTAGAAAGCATG 2037  
DB 1978 CATGATGCTCTGTGCAAAATTAATGAGAGAGTGGAGTGGAAAACTAGAAAGCATG 2037  
QY 1981 CAGATGCTCTGTGCAAAATTAATGAGAGAGTGGAGTGGAAAACTAGAAAGCATG 2040  
DB 1981 CAGATGCTCTGTGCAAAATTAATGAGAGAGTGGAGTGGAAAACTAGAAAGCATG 2040  
QY 2038 AAAGACCTGSCCACTTTAAAGAGAAAGAGCTACCTTCAGAGATTTTGTGTCATTA 2094  
DB 2038 AAAGACCTGSCCACTTTAAAGAGAAAGAGCTACCTTCAGAGATTTTGTGTCATTA 2094  
QY 2041 AAAACCTGSCCACTTTAAAGAGAAAGAGCTACCTTCAGAGATTTTGTGTCATTA 2097  
DB 2041 AAAACCTGSCCACTTTAAAGAGAAAGAGCTACCTTCAGAGATTTTGTGTCATTA 2097

RESULT 5  
US-09-371-347-43  
; Sequence 43, Application US/09371347  
; Publication No. US20030082676A1  
; GENERAL INFORMATION:  
; APPLICANT: Roy A. Gravel et al.  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE REDUCTASE:  
; TITLE OF INVENTION: CLONING, AND METHODS FOR EVALUATING RISK OF NEURAL TUBE  
; TITLE OF INVENTION: DEFECTS, CARDIOVASCULAR DISEASE, AND CANCER  
; FILE REFERENCE: 50004/003003  
; CURRENT APPLICATION NUMBER: US/09/371,347

; CURRENT FILING DATE: 1999-08-10  
; PRIOR APPLICATION NUMBER: 60/071,622  
; PRIOR FILING DATE: 1998-01-16  
; PRIOR APPLICATION NUMBER: 09/232,028  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: FastSeq For Windows Version 4.0  
; SEQ ID NO 43  
; LENGTH: 2097  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-371-347-43

Query Match 86.1%; Score 1803; DB 10; Length 2097;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2093; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 1 ATGAGAGAGTTTCTGTACTATATGCTATACACAGAGGACAGGCAAGGCCATCGAGAA 60  
DB 1 ATGAGAGAGTTTCTGTACTATATGCTATACACAGAGGACAGGCAAGGCCATCGAGAA 60  
QY 61 GAAATGTGAGCAGAGCTGTGATCATGATTTTCTGAGATCTTCACTGATTAATGAA 120  
DB 61 GAAATGTGAGCAGAGCTGTGATCATGATTTTCTGAGATCTTCACTGATTAATGAA 120  
QY 121 TCCGATATGATGACCTTAAACCGAAGAGCTCTTGTGTGTGTTCTACACG 180  
DB 121 TCCGATATGATGACCTTAAACCGAAGAGCTCTTGTGTGTGTTCTACACG 180  
QY 181 GGCACCGGAGACCCACCCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCA 240  
DB 181 GGCACCGGAGACCCACCCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCA 240  
QY 241 CTGCGGTTGATTTCTTGTCTCACTGCGGTAATGAGTACTGGGCTCGGTGATTCAGAA 300  
DB 241 CTGCGGTTGATTTCTTGTCTCACTGCGGTAATGAGTACTGGGCTCGGTGATTCAGAA 300  
QY 301 TACACCTACTTTTGCATGAGGAGGAGATTAATGATTAACGACTTCAAGACTTGAACC 360  
DB 301 TACACCTACTTTTGCATGAGGAGGAGATTAATGATTAACGACTTCAAGACTTGAACC 360  
QY 361 CGGCAATTTCTATGACACTGAGACATGAGATGACTGTGAGTTTAAACCTGTGGTTGAG 420  
DB 361 CGGCAATTTCTATGACACTGAGACATGAGATGACTGTGAGTTTAAACCTGTGGTTGAG 420  
QY 421 CCGTGAATGCTGGAATCTGCGCAGCCCTCAGAAACATTTTATGCTCAAGAGAGCA 480  
DB 421 CCGTGAATGCTGGAATCTGCGCAGCCCTCAGAAACATTTTATGCTCAAGAGAGCA 480  
QY 481 GAGAGATATGAGGAGCTCCCGGTGATCACTGATCTTGAAGACAGACTTTGTG 540  
DB 481 GAGAGATATGAGGAGCTCCCGGTGATCACTGATCTTGAAGACAGACTTTGTG 540  
QY 541 AAGTCAGAGCTGTACATTAATGATCTCAAGTGCAGCTTGTGATTTCAATGATTTAGGA 600  
DB 541 AAGTCAGAGCTGTACATTAATGATCTCAAGTGCAGCTTGTGATTTCAATGATTTAGGA 600  
QY 601 AGAAGATTTCTGAGGTTTGAAGCAAAATGACAGTAAAGCAACCAATTCATGTTGTA 660  
DB 601 AGAAGATTTCTGAGGTTTGAAGCAAAATGACAGTAAAGCAACCAATTCATGTTGTA 660  
QY 661 ATTGAAGACTTTGAGTCTTCACTTACCCGTTGAGTACCCCACTCTCAAGCTCTCTG 720  
DB 661 ATTGAAGACTTTGAGTCTTCACTTACCCGTTGAGTACCCCACTCTCAAGCTCTCTG 720  
QY 721 AATATTCCTGTTTACCCCAAGAAATTTTACAGATCACTGAGAGAGTCTTTGGCCAG 780  
DB 721 AATATTCCTGTTTACCCCAAGAAATTTTACAGATCACTGAGAGAGTCTTTGGCCAG 780  
QY 781 GAGAGAGCAAGATATCTGATCTTCAAGAGATCCAGATTTTCAAGTGCATTTCAAG 840  
DB 781 GAGAGAGCAAGATATCTGATCTTCAAGAGATCCAGATTTTCAAGTGCATTTCAAG 840

QY 841 GCAGTTCACCTTACTAGATGATGTCATTAACCACTCTGTGTAGATTTGACAT 900  
 DB 841 GCAGTTCACCTTACTAGATGATGTCATTAACCACTCTGTGTAGATTTGACAT 900  
 QY 901 TCAATACAGACTTTTCTATCAGCTGAGATGCTTGAAGGTATGCTTAAAGT 960  
 DB 901 TCAATACAGACTTTTCTATCAGCTGAGATGCTTGAAGGTATGCTTAAAGT 960  
 QY 961 GATTCGAGGTACAAAGCCCTACCCAAAGCTGAGCTTGAAGATTAAGAGAGAGCTGC 1020  
 DB 961 GATTCGAGGTACAAAGCCCTACCCAAAGCTGAGCTTGAAGATTAAGAGAGAGCTGC 1020  
 QY 1021 GTCTTTTGAATAAAGGACAGACAAAGAAAGAGAGCTTACCCAGCATATA 1080  
 DB 1021 GTCTTTTGAATAAAGGACAGACAAAGAAAGAGAGCTTACCCAGCATATA 1080  
 QY 1081 CCTGCGGAGATGTTCTCTCAGATTCATTTTACCTGTGTCTTGAATCCGAGCATTTCT 1140  
 DB 1081 CCTGCGGAGATGTTCTCTCAGATTCATTTTACCTGTGTCTTGAATCCGAGCATTTCT 1140  
 QY 1141 AAAAAGGCAATTTTGGAGAGCCCTTGGAGCTATACAGAGAGAGCTGAAAAGCCGAG 1200  
 DB 1141 AAAAAGGCAATTTTGGAGAGCCCTTGGAGCTATACAGAGAGAGCTGAAAAGCCGAG 1200  
 QY 1201 CTACAGAGCTGTGACGTAAACAAAGGGGAGCCGATTTATAGCCGTTTGTACAGATGCC 1260  
 DB 1201 CTACAGAGCTGTGACGTAAACAAAGGGGAGCCGATTTATAGCCGTTTGTACAGATGCC 1260  
 QY 1261 TGTGCTGTGTGTGATCTCTCTGCTTCTCTTCTTTCGAGCCAGCAGCTCAGTCTTC 1320  
 DB 1261 TGTGCTGTGTGTGATCTCTCTGCTTCTCTTCTTTCGAGCCAGCAGCTCAGTCTTC 1320  
 QY 1321 CTGCTCGAATCTTCTTAACTTCAACCCAGACATTTCTGTGTCAAGCTCAATTTA 1380  
 DB 1321 CTGCTCGAATCTTCTTAACTTCAACCCAGACATTTCTGTGTCAAGCTCAATTTA 1380  
 QY 1381 TTTACCCAGAAAGCTTCATTTTGTCTTCAACATTTGGAATTTCTGTCTACTGACACA 1440  
 DB 1381 TTTACCCAGAAAGCTTCATTTTGTCTTCAACATTTGGAATTTCTGTCTACTGACACA 1440  
 QY 1441 ACAGAGGTTCTGCGAAGGAGATGTATACAGCTGAGCTGTGTGTGTGTGTGTGTGTGT 1500  
 DB 1441 ACAGAGGTTCTGCGAAGGAGATGTATACAGCTGAGCTGTGTGTGTGTGTGTGTGTGT 1500  
 QY 1501 CTTCAGCCAAACATATCATGATCTCCATGAAGAGAGGGGAAAGCCCTGCTCTTAAGATA 1560  
 DB 1501 CTTCAGCCAAACATATCATGATCTCCATGAAGAGAGGGGAAAGCCCTGCTCTTAAGATA 1560  
 QY 1561 TCCATCTCTCTGAGCAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
 DB 1561 TCCATCTCTCTGAGCAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
 QY 1621 ATATAGTGTGATTCAGAAACCGGACATAGCCCGTTTATTTGAGTTCTTACAACTAAGAG 1680  
 DB 1621 ATATAGTGTGATTCAGAAACCGGACATAGCCCGTTTATTTGAGTTCTTACAACTAAGAG 1680  
 QY 1681 AAATCTCAAGAAACAAACCCAGATGAAATTTTGAAGCAATGTG---GTTTTTGGCTGC 1737  
 DB 1681 AAATCTCAAGAAACAAACCCAGATGAAATTTTGAAGCAATGTG---GTTTTTGGCTGC 1737  
 QY 1738 AGGCAATAGGATAGGATTTATCTATTCAAGAAAGAGCTGACATTTCTTAAAGAGAG 1797  
 DB 1738 AGGCAATAGGATAGGATTTATCTATTCAAGAAAGAGCTGACATTTCTTAAAGAGAG 1797  
 QY 1801 ATCTTAATCATTTAAGGTTTCTTCTCAAGAGATGCTCTGTGTGGAGAGAGAGAGCC 1860  
 DB 1801 ATCTTAATCATTTAAGGTTTCTTCTCAAGAGATGCTCTGTGTGGAGAGAGAGAGCC 1860  
 QY 1858 CCAGCAAGATATTAACAAGCAATTCAGCTTCAATGAGAGAGAGAGAGAGAGAGAGAG 1917  
 DB 1858 CCAGCAAGATATTAACAAGCAATTCAGCTTCAATGAGAGAGAGAGAGAGAGAGAGAG 1917  
 QY 1861 CCAGCAAGATATTAACAAGCAATTCAGCTTCAATGAGAGAGAGAGAGAGAGAGAG 1920  
 DB 1861 CCAGCAAGATATTAACAAGCAATTCAGCTTCAATGAGAGAGAGAGAGAGAGAGAG 1920  
 QY 1918 CTCAG 1977  
 DB 1918 CTCAG 1977

DB 1921 CTCAG 1980  
 QY 1978 CATGATGCCCTGTGCAATTAATTAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2037  
 DB 1981 CATGATGCCCTGTGCAATTAATTAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2040  
 QY 2038 AAAACCTGGCAGCTTTAAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2094  
 DB 2041 AAAACCTGGCAGCTTTAAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2097

RESULT 6  
 US-09-371-347-47  
 ; Sequence 47, Application US/09371347  
 ; Publication No. US20030082676A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Roy A. Gravel et al.  
 ; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE REDUCTASE;  
 ; TITLE OF INVENTION: CLONING, AND METHODS FOR EVALUATING RISK OF NEURAL TUBE  
 ; TITLE OF INVENTION: DEFECTS CARDIOVASCULAR DISEASE, AND CANCER  
 ; FILE REFERENCE: 50004/003003  
 ; CURRENT APPLICATION NUMBER: US/09/371,347  
 ; PRIOR FILING DATE: 1999-08-10  
 ; PRIOR APPLICATION NUMBER: 60/071,622  
 ; PRIOR FILING DATE: 1998-01-16  
 ; PRIOR APPLICATION NUMBER: 09/232,028  
 ; PRIOR FILING DATE: 1999-01-15  
 ; NUMBER OF SEQ ID NOS: 51  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 47  
 ; LENGTH: 2093  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-371-347-47

Query Match 80.1%; Score 1677; DB 10; Length 2093;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1677; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGAGGTTCTCTTACTATATGCTATACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60  
 DB 1 ATGAGAGGTTCTCTTACTATATGCTATACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60  
 QY 61 GAATGTGTGACCAAGCTGTGTGATCATGATTTCTGAGATTTCACTATTAAGATA 120  
 DB 61 GAATGTGTGACCAAGCTGTGTGATCATGATTTCTGAGATTTCACTATTAAGATA 120  
 QY 121 TCCGATTAAGATGACCTTAATAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
 DB 121 TCCGATTAAGATGACCTTAATAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
 QY 181 GGCACCGAG 240  
 DB 181 GGCACCGAG 240  
 QY 241 CTGCGGATGATTTCTTGTCTCACTGCGGATATGAGTTCTGGTCTCGGATTTCAAGAA 300  
 DB 241 CTGCGGATGATTTCTTGTCTCACTGCGGATATGAGTTCTGGTCTCGGATTTCAAGAA 300  
 QY 301 TACACCTACTTTTGTGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
 DB 301 TACACCTACTTTTGTGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
 QY 361 CCGCATTTCTATGACACTGAG 420  
 DB 361 CCGCATTTCTATGACACTGAG 420  
 QY 421 CCGTGAATGCTGAGATCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
 DB 421 CCGTGAATGCTGAGATCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
 QY 481 GAGAGATTAAG 540  
 DB 481 GAGAGATTAAG 540

481 GAGAGGTAAGTGGCGACCTCCGGTGGCACTCCTGACCTTGGAGAGCAAGCTTGTG 540  
541 AAGTCAGAGCTGCTACACATTTGAATCTCAAGTCGAGCTTCTGAGATTGATTCAGGA 600  
541 AAGTCAGAGCTGCTACACATTTGAATCTCAAGTCGAGCTTCTGAGATTGATTCAGGA 600  
601 AGAAGGATTTCTGAGCTTTGAAAGCAAAATGCAAGTGAACGCAACCAATCCAAATGTTGA 660  
601 AGAAGGATTTCTGAGCTTTGAAAGCAAAATGCAAGTGAACGCAACCAATCCAAATGTTGA 660  
661 ATTGAAGCTTTGAGCTCTCACTTACCGTGGTACCCCACTCTCAAGAGCTCTGTG 720  
661 ATTGAAGCTTTGAGCTCTCACTTACCGTGGTACCCCACTCTCAAGAGCTCTGTG 720  
721 AATATTCCTGTTTACCCCGAATATTTAAGGTATCATCTGACGAGAGTCTTGGCCAG 780  
721 AATATTCCTGTTTACCCCGAATATTTAAGGTATCATCTGACGAGAGTCTTGGCCAG 780  
781 GAGGAAAGCCAAAGTATCTGTGATCTTCAAGCATTCAGATCTTTTCAAGTCCAAATTCAG 840  
781 GAGGAAAGCCAAAGTATCTGTGATCTTCAAGCATTCAGATCTTTTCAAGTCCAAATTCAG 840  
841 GCAGTTCAACTTCTACGAATGATGCAATAAAACACTGCTGGTGAATTTGACATT 900  
841 GCAGTTCAACTTCTACGAATGATGCAATAAAACACTGCTGGTGAATTTGACATT 900  
901 TCAATATCAGACTTTTCTATTCAGCCCTGAGATGCTTCAAGCTGATCTGCTTAACTG 960  
901 TCAATATCAGACTTTTCTATTCAGCCCTGAGATGCTTCAAGCTGATCTGCTTAACTG 960  
961 GATTCGAGGTACAAAGCTTCTCAAGCTGAGATGCTTCAAGCTGAGATGCTTCAAGCTG 1020  
961 GATTCGAGGTACAAAGCTTCTCAAGCTGAGATGCTTCAAGCTGAGATGCTTCAAGCTG 1020  
1021 GTCCTTTTGAATTAAGGAGACACAAAGAAAGAGAGTACCTTACCCCAAGCATTA 1080  
1021 GTCCTTTTGAATTAAGGAGACACAAAGAAAGAGAGTACCTTACCCCAAGCATTA 1080  
1081 CCGTGGGAGATGTTCTCTCAAGTCAATTTTACCTGCTGCTTGAATTCGAGCAATTTCT 1140  
1081 CCGTGGGAGATGTTCTCTCAAGTCAATTTTACCTGCTGCTTGAATTCGAGCAATTTCT 1140  
1141 AAAAAGGCAATTTTGGAGAGCCCTTGTGACTATACAGTGAAGTGTGAAAAGCCGAG 1200  
1141 AAAAAGGCAATTTTGGAGAGCCCTTGTGACTATACAGTGAAGTGTGAAAAGCCGAG 1200  
1201 CTAAGAGAGCTGTGCAATTAACAAGGGGAGCCGATTAATAGCCGCTTGTGAGAGATGCC 1260  
1201 CTAAGAGAGCTGTGCAATTAACAAGGGGAGCCGATTAATAGCCGCTTGTGAGAGATGCC 1260  
1261 TGTGCTGCTGTTGATCTCTCTCGCTTCCCTTCTTGCAGCCCACTCAAGTCTC 1320  
1261 TGTGCTGCTGTTGATCTCTCTCGCTTCCCTTCTTGCAGCCCACTCAAGTCTC 1320  
1321 CTGCTGCAATCTTCTCTAACTTCAACCAAGACATATTCGTGTGAGAGCTCAAGTTTA 1380  
1321 CTGCTGCAATCTTCTCTAACTTCAACCAAGACATATTCGTGTGAGAGCTCAAGTTTA 1380  
1381 TTTCAACCCAGAAAGCTCAATTTTGTCTCAACATGTGGAATTTCTGTCTACGCGACA 1440  
1381 TTTCAACCCAGAAAGCTCAATTTTGTCTCAACATGTGGAATTTCTGTCTACGCGACA 1440  
1441 ACAGAGGTTCTGCGAAGGAGATATGTAAGCTGCTGCGCTTGTGTTGCTTCAAGTT 1500  
1441 ACAGAGGTTCTGCGAAGGAGATATGTAAGCTGCTGCGCTTGTGTTGCTTCAAGTT 1500  
1501 CTTCAAGCAAAATCATCTGATCCCATGAAGACAGCGGAAAGCCCTGCTCTTAAGATA 1560  
1501 CTTCAAGCAAAATCATCTGATCCCATGAAGACAGCGGAAAGCCCTGCTCTTAAGATA 1560  
1561 TCCATCTCTCTGGAACAACAATTTCTTTCATCTTACAGATGACCCCTCAATCCCATC 1620  
1561 TCCATCTCTCTGGAACAACAATTTCTTTCATCTTACAGATGACCCCTCAATCCCATC 1620

1561 TCCATCTCTCTGGAACAACAATTTCTTTCATCTTACAGATGACCCCTCAATCCCATC 1620  
1621 ATATGATGAGTCTCAGAAACCGGACATAGCCCGTTTATTTGGTTCTTACAAATAGA 1677  
1621 ATATGATGAGTCTCAGAAACCGGACATAGCCCGTTTATTTGGTTCTTACAAATAGA 1677  
RESULT 7  
US-10-741-600-692  
; Sequence 692, Application US/10741600  
; Publication No. US20050026169A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; FILE REFERENCE: CL001499  
; CURRENT APPLICATION NUMBER: US/10/741,600  
; NUMBER OF SEQ ID NOS: 73997  
; SOFTWARE: PatSeq for Windows Version 4.0  
; SEQ ID NO 692  
; LENGTH: 3256  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-741-600-692  
Query Match 42.8%; Score 896; DB 21; Length 3256;  
Best Local Similarity 99.1%; Pred. No. 0;  
Matches 1646; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
67 TGTGAGCAAGCTGTGTGATCATGATTTTCTGAGATCTTCACTGTATTTAGTAATCCGAT 126  
160 TGTGAGCAAGCTGTGTGATCATGATTTTCTGAGATCTTCACTGTATTTAGTAATCCGAT 219  
127 AAGTATGACCTTAAACCGAAACAGCTCTCTGTTGTTGTTGTTTCTACAGGGGAC 186  
220 AAGTATGACCTTAAACCGAAACAGCTCTCTGTTGTTGTTGTTTCTACAGGGGAC 219  
187 GGAAGCCACCGGACACAGCCCGGAAATTTGTAAGAAATACAGAACCAACACTGCCG 246  
280 GGAAGCCACCGGACACAGCCCGGAAATTTGTAAGAAATACAGAACCAACACTGCCG 339  
247 GTTGATTTCTTTGCTCACTGCGGTATGTTTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 306  
340 GTTGATTTCTTTGCTCACTGCGGTATGTTTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 399  
307 TACTTTGCAATGGGGGAAATTTGATTAACGACTTCAAGAGCTTGAAGCCCGCAT 366  
400 TACTTTGCAATGGGGGAAATTTGATTAACGACTTCAAGAGCTTGAAGCCCGCAT 459  
367 TTCTATGACATGACATGATGATCTGTGATGTTTGAACCTTGTGTTGAGCCGTGG 426  
460 TTCTATGACATGACATGATGATCTGTGATGTTTGAACCTTGTGTTGAGCCGTGG 519  
427 ATTGCTGACTCTGCGACCCCTCAGAAAGCATTTTATGCTCAAGAGAGAGAGAG 486  
520 ATTGCTGACTCTGCGACCCCTCAGAAAGCATTTTATGCTCAAGAGAGAGAGAGAG 579  
487 ATTAAGTGGGCACTCCCGGTGGATCACTGATCTCTTGAAGAGACAGCTTGAAGTGA 546  
580 ATTAAGTGGGCACTCCCGGTGGATCACTGATCTCTTGAAGAGAGAGAGAGAGAG 639  
547 GAGCTGCTACATTAATCTCAAGTCAAGCTTCTGAGATTTGAGAGAGAGAG 606  
640 GAGCTGCTACATTAATCTCAAGTCAAGCTTCTGAGATTTGAGAGAGAGAGAG 699  
607 GATTTGAGGTTTGAAGCAAAATGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 666  
700 GATTTGAGGTTTGAAGCAAAATGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 759  
667 GACTTGAAGCTTCACTTACCCGTTGATACCCCACTCTCAAGAGCTTCTGAATAT 726  
760 GACTTGAAGCTTCACTTACCCGTTGATACCCCACTCTCAAGAGCTTCTGAATAT 819

QY 727 CCGGTTTACCCCAAGATATTTAGAGTACATCTGACAGAGTCTCTGGCCAGAGGAA 786  
DB 820 CCGGTTTACCCCAAGATATTTAGAGTACATCTGACAGAGTCTCTGGCCAGAGGAA 879  
QY 787 AGCCAAATCTGTGATCTCAGCAGATCCAGTTTCAAGTGCATTTCAAGGAGTT 846  
DB 880 AGCCAAATCTGTGATCTCAGCAGATCCAGTTTCAAGTGCATTTCAAGGAGTT 939  
QY 847 CAATTTACTACGAATATGATCAATTAACCACTCTGCTGTAGAAATTTGACATTTCAAT 906  
DB 940 CAATTTACTACGAATATGATCAATTAACCACTCTGCTGTAGAAATTTGACATTTCAAT 999  
QY 907 ACAGACTTTTCTATGAGCTGAGATGCTTCAAGGATCTGAGCTTCAAGTATTC 966  
DB 1000 ACAGACTTTTCTATGAGCTGAGATGCTTCAAGGATCTGAGCTTCAAGTATTC 1059  
QY 967 GAGTACAAAGCTTCTCAAGATCTGAGCTTGAAGATTAAGAGAGCACTGCTCTT 1026  
DB 1060 GAGTACAAAGCTTCTCAAGATCTGAGCTTGAAGATTAAGAGAGCACTGCTCTT 1119  
QY 1027 TTGAATTAAGGAGAGCAACAAAGAAAGAGCTACCTTACCCAGCATATCTGCG 1086  
DB 1120 TTGAATTAAGGAGAGCAACAAAGAAAGAGCTACCTTACCCAGCATATCTGCG 1179  
QY 1087 GAGTGTCTCTCAGTTCAATTTTACCTGCTGTCTGAATCCGAGCAATTCCTAATAAG 1146  
DB 1180 GAGTGTCTCTCAGTTCAATTTTACCTGCTGTCTGAATCCGAGCAATTCCTAATAAG 1239  
QY 1147 GCATTTTTCGAGCCCTTGTGATCTATACAGTACGATGCTGAAGAGGCAAGCTACAG 1206  
DB 1240 GCATTTTTCGAGCCCTTGTGATCTATACAGTACGATGCTGAAGAGGCAAGCTACAG 1299  
QY 1207 GAGCTGTGATTAACAAGAGGAGCGAGCTATTAAGCCGCTTTGTACAGATGCTGTGCC 1266  
DB 1300 GAGCTGTGATTAACAAGAGGAGCGAGCTATTAAGCCGCTTTGTGTGTGTGTGTGTGCC 1359  
QY 1267 TGTCTGT 1326  
DB 1360 TGTCTGT 1419  
QY 1327 GAACATCTTCTTAATCTTAACCCAGACATATTCGTGTGCAAGCTCAAGTTTATTTAC 1386  
DB 1420 GAACATCTTCTTAATCTTAACCCAGACATATTCGTGTGCAAGCTCAAGTTTATTTAC 1479  
QY 1387 CCAAGAAAGCTCCTATTTGTCTTCAATTTGTGAATTTCTGTCTACCTGCAACAGAG 1446  
DB 1480 CCAAGAAAGCTCCTATTTGTCTTCAATTTGTGAATTTCTGTCTACCTGCAACAGAG 1539  
QY 1447 GTTCTGCGAAGGAGATATGACAGGCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGT 1506  
DB 1540 GTTCTGCGAAGGAGATATGACAGGCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGT 1599  
QY 1507 CCAAAATATCATGATCCCATGAAGACAGGAGGAAAGCCCTGCTCTTAAGATATCATC 1566  
DB 1600 CCAAAATATCATGATCCCATGAAGACAGGAGGAAAGCCCTGCTCTTAAGATATCATC 1659  
QY 1567 TCTCTGCAACAAATTTCTTCACTTACAGATGACCTTCAATCCCATCATPATG 1626  
DB 1660 TCTCTGCAACAAATTTCTTCACTTACAGATGACCTTCAATCCCATCATPATG 1719  
QY 1627 GTGGGTTCAGAAACCGGATAGCCGCTTATTTGGGTCTCTAACAATAGAGAAATCTC 1686  
DB 1720 GTGGGTTCAGAAACCGGATAGCCGCTTATTTGGGTCTCTAACAATAGAGAAATCTC 1779  
QY 1687 CAGAACAACACCCAGATGGAATTTTGGAGCATGTGTT 1727  
DB 1780 CAGAACAACACCCAGATGGAATTTTGGAGCATGTGTT 1820

RESULT 8  
US-10-741-600-693  
; Sequence 693, Application US/10741600

Publication No. US2005002619A1  
GENERAL INFORMATION:  
APPLICANT: CARILL, Michele et al.  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
FILE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001499  
CURRENT APPLICATION NUMBER: US/10/741,600  
NUMBER OF SEQ ID NOS: 2003-12-22  
NUMBER OF SEQ ID NOS: 73997  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 693  
LENGTH: 3274  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-741-600-693  
Query Match 42.8%; Score 896; DB 21; Length 3274;  
Best Local Similarity 99.1%; Pred. No. 0;  
Matches 1646; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
QY 67 TGTGAGCAAGCTGTGTACATGATTTTCTGAGATCTTCACTGATTTAGTGAATCCGAT 126  
DB 178 TGTGAGCAAGCTGTGTACATGATTTTCTGAGATCTTCACTGATTTAGTGAATCCGAT 237  
QY 127 AAGTATGACCTTAATAACCGAAACAGCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 186  
DB 238 AAGTATGACCTTAATAACCGAAACAGCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 297  
QY 187 GAGAGCCCAACCCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACACTGCCG 246  
DB 298 GAGAGCCCAACCCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACACTGCCG 357  
QY 247 GTTGAATTTCTTGTCTCACTCGGTGTATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 306  
DB 358 GTTGAATTTCTTGTCTCACTCGGTGTATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 417  
QY 307 TACTTTTGCATGAGGAGGAGAAATTTGATTAACGACTTCAAGAGCTTGAAGCCCGCAT 366  
DB 418 TACTTTTGCATGAGGAGGAGAAATTTGATTAACGACTTCAAGAGCTTGAAGCCCGCAT 477  
QY 367 TTCTATGACACTGACATGACAGATGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 426  
DB 478 TTCTATGACACTGACATGACAGATGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 537  
QY 427 ATTGCTGACCTGTGCTGACCCCTCAGAAAGCAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGT 486  
DB 538 ATTGCTGACCTGTGCTGACCCCTCAGAAAGCAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGT 597  
QY 487 ATTAAGTGGGCACTCCCGGTGGATCACTGTGATCTTGTAGAGACAGACCTTGTGAAGTCA 546  
DB 598 ATTAAGTGGGCACTCCCGGTGGATCACTGTGATCTTGTAGAGACAGACCTTGTGAAGTCA 657  
QY 547 GAGCTGCTACATGATTAATCTCAAGTGAAGCTTGTGATTTGATTTCAAGAAAGAG 606  
DB 658 GAGCTGCTACATGATTAATCTCAAGTGAAGCTTGTGATTTGATTTCAAGAAAGAG 717  
QY 607 GATTTGAGGTTTGTGAAGAAATGTAGTGAACAGAACCAATCCATGTTGTATTTGA 666  
DB 718 GATTTGAGGTTTGTGAAGAAATGTAGTGAACAGAACCAATCCATGTTGTATTTGA 777  
QY 667 GACTTGAAGTCTCACTTACCCGTTGAGTACCCCACTCTCAAGCTCTCTGAATATT 726  
DB 778 GACTTGAAGTCTCACTTACCCGTTGAGTACCCCACTCTCAAGCTCTCTGAATATT 837  
QY 727 CCGGTTTACCCCAAGATATTTAAGATATCTGACAGAGTCTTGGCCAGAGGAA 786  
DB 838 CCGGTTTACCCCAAGATATTTAAGATATCTGACAGAGTCTTGGCCAGAGGAA 897  
QY 787 AGCCAAATCTGTGATCTCAGCAGATCCAGTTTCAAGTGCATTTCAAGGAGTT 846  
DB 898 AGCCAAATCTGTGATCTCAGCAGATCCAGTTTCAAGTGCATTTCAAGGAGTT 957  
QY 847 CAATTTACTACGAATATGATCAATTAACCACTCTGCTGTAGAAATTTGACATTTCAAT 906

```
Db 958 CAACCTTACGAGTATGATCCATTAACCACTCTGCTGATGAAATTTGACATTTCAANT 1017
Qy 907 ACAGACTTTTCTATACAGCTGAGATGCTTTCAGCGTGATCTGCCCTAACAGTATTC 966
Db 1018 ACAGACTTTTCTATGAGCTGAGATGCTTTCAGCGTGATCTGCCCTAACAGTATTC 1077
Qy 967 GAGGTCAAAAGCTTATCCCAAGATGCTGAGCTTGAAGATTAAGAGACCTGCTCTT 1026
Db 1078 GAGGTCAAAAGCTTATCCCAAGATGCTGAGCTTGAAGATTAAGAGACCTGCTCTT 1137
Qy 1027 TTGAAATATAAGGACAGACAAAGAAAGAGACTACCTTACCCCGCATATACCTGCG 1086
Db 1138 TTGAAATATAAGGACAGACAAAGAAAGAGACTACCTTACCCCGCATATACCTGCG 1197
Qy 1087 GAGTGTCTCTCAGATTTATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1146
Db 1198 GAGTGTCTCTCAGATTTATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1257
Qy 1147 GCATTTTGGAGACCTTGTGAGCTATACAGTGAACAGTGTGTAAGAGAGAGAGAGAG 1206
Db 1258 GCATTTTGGAGACCTTGTGAGCTATACAGTGAACAGTGTGTAAGAGAGAGAGAGAG 1317
Qy 1207 GAGCTGTGAGTAAACAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1266
Db 1318 GAGCTGTGAGTAAACAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1377
Qy 1267 TGCCTTGTGATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1326
Db 1378 TGCCTTGTGATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1437
Qy 1327 GAACATCTTCTTAACTTCAACCCAGACCATATTCGCTGAGAGAGAGAGAGAGAGAGAG 1386
Db 1438 GAACATCTTCTTAACTTCAACCCAGACCATATTCGCTGAGAGAGAGAGAGAGAGAGAG 1497
Qy 1387 CCAGAAAGCTCATTTTGTCTTCAACATGTGGAATTTCTGCTGCTGCTGCTGCTGCTGCTG 1446
Db 1498 CCAGAAAGCTCATTTTGTCTTCAACATGTGGAATTTCTGCTGCTGCTGCTGCTGCTGCTG 1557
Qy 1447 GTTCTGGAAGAGAGATGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1506
Db 1558 GTTCTGGAAGAGAGATGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1617
Qy 1507 CCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1566
Db 1618 CCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1677
Qy 1567 TCTCTGCAACCAAAATTTCTTCACTTACAGATGACCCCTCAATCCCATCATATG 1626
Db 1678 TCTCTGCAACCAAAATTTCTTCACTTACAGATGACCCCTCAATCCCATCATATG 1737
Qy 1627 GTGGGTTCAGAAACCGGATAGCCCGTTTATTTGAGTTCTTACACATAGAGAAATC 1686
Db 1738 GTGGGTTCAGAAACCGGATAGCCCGTTTATTTGAGTTCTTACACATAGAGAAATC 1797
Qy 1687 CAAGAACAAACCAAGATGGAATTTTGGAGCATGTGTT 1727
Db 1798 CAAGAACAAACCAAGATGGAATTTTGGAGCATGTGTT 1838

RESULT 9
US-10-029-386-6369
; Sequence 6369, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
```

```
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 6369
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC008727.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.45
; OTHER INFORMATION: NT HIT: AF121205.1, EVALU0 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: Q9UPW0, EVALU0 3.30e+00
; OTHER INFORMATION: EST_HUMAN HIT: AV132586.1, EVALU0 0.00e+00
US-10-029-386-6369

Query Match 15.8%; Score 330; DB 16; Length 591;
Best Local Similarity 99.7%; Pred. No. 1.7e-169;
Matches 380; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 401 GTTTGAACCTTGATGTTGAGCGGTGATGCTGGAATGCTGGCCAGCCCTCAGAAAGCAT 460
Db 38 GTTTGAACCTTGATGTTGAGCGGTGATGCTGGAATGCTGGCCAGCCCTCAGAAAGCAT 97
Qy 461 TTAGGTCAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 520
Db 98 TTAGGTCAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 157
Qy 521 CCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 580
Db 158 CCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 217
Qy 581 TGAGATTCATGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 640
Db 218 TGAGATTCATGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 277
Qy 641 GCACCAATCCATGTTGATTAATGAAGCTTTGAGTCTCACTTACCCGTTGCTGCTGCTG 700
Db 278 GCACCAATCCATGTTGATTAATGAAGCTTTGAGTCTCACTTACCCGTTGCTGCTGCTG 337
Qy 701 CACTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 760
Db 338 CACTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 397
Qy 761 TGAGAGAGTCTTGCCAGG 781
Db 398 TGAGAGAGTCTTGCCAGG 418

RESULT 10
US-10-029-386-20100
; Sequence 20100, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 20100
; LENGTH: 379
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC008727.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.45
; OTHER INFORMATION: NT HIT: G114729757, EVALU0 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: Q9UPW0, EVALU0 1.80e+00
; OTHER INFORMATION: EST_HUMAN HIT: AV132586.1, EVALU0 0.00e+00
US-10-029-386-20100
```

Query Match 15.7%; Score 328; DB 16; Length 379;  
Best Local Similarity 99.7%; Pred. No. 26-168;  
Matches 378; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 402 TTAAAGCTTGTGTGAGCCCTGAGATTCCTGAGCTTCGCGAGCCCTCAGAAAGCATTT 461  
DB 1 TTAAAGCTTGTGTGAGCCCTGAGATTCCTGAGCTTCGCGAGCCCTCAGAAAGCATTT 60  
QY 462 TAGGTCAAGCAGAGCAGAGAGATTAAGTGGCCCACTCCCGGTGGCATTCCTGATC 521  
DB 61 TAGGTCAAGCAGAGCAGAGAGATTAAGTGGCCCACTCCCGGTGGCATTCCTGATC 120  
QY 522 CTGAGAGCAGACCTTGTGAAGTCAAGAGCTGCTACACATTTAATCTCAAGTCACTT 581  
DB 121 CTCGAGAGCAGACCTTGTGAAGTCAAGAGCTGCTACACATTTAATCTCAAGTCACTT 180  
QY 582 GAGATTGATGATTCAGAGAAAGAGATTCAGAGTTTGAAGCAAAATGCAGTGAACG 641  
DB 181 GAGATTGATGATTCAGAGAAAGAGATTCAGAGTTTGAAGCAAAATGCAGTGAACG 240  
QY 642 CAACCAATCCAACTGTGTAATTTGAAGACTTTGAGTCTTCACTTACCCGTTCCGTAACCC 701  
DB 241 CAACCAATCCAACTGTGTAATTTGAAGACTTTGAGTCTTCACTTACCCGTTCCGTAACCC 300  
QY 702 ACTCTCAAGCCCTCTGAATATTCCTGCTTACCCCGAATATTTACAGTACATCT 761  
DB 301 ACTCTCAAGCCCTCTGAATATTCCTGCTTACCCCGAATATTTACAGTACATCT 360  
QY 762 GCAGAGTCTCTTGGCCAG 780  
DB 361 GCAGAGTCTCTTGGCCAG 379

RESULT 11  
US-10-029-386-1735  
Sequence 1735, Application US/10029386  
Publication No. US20030194704A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Hanzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
FILE REFERENCE: AEOMICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 34288  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 1735  
LENGTH: 591  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC021609.3  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.4  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.8  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.2  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.2  
OTHER INFORMATION: NT HIT: AF121205.1, EVALUE 0.00e+00  
OTHER INFORMATION: EST HUMAN HIT: AU132586.1, EVALUE 0.00e+00  
OTHER INFORMATION: SWISSPROT HIT: Q9UPW0, EVALUE 3.30e+00  
US-10-029-386-1735

Query Match 13.3%; Score 279; DB 16; Length 591;  
Best Local Similarity 99.5%; Pred. No. 1.6e-141;  
Matches 379; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 401 GTTAAAGCTTGTGTGAGCCCTGAGATTCCTGAGCTTCGCGAGCCCTCAGAAAGCATTT 460  
DB 38 GTTAAAGCTTGTGTGAGCCCTGAGATTCCTGAGCTTCGCGAGCCCTCAGAAAGCATTT 97

QY 461 TTAGGTCAAGCAGAGCAGAGAGATTAAGTGGCCCACTCCCGGTGGCATTCCTGATC 520  
DB 98 TTAGGTCAAGCAGAGCAGAGAGATTAAGTGGCCCACTCCCGGTGGCATTCCTGATC 157  
QY 521 CCTGAGAGCAGACCTTGTGAAGTCAAGAGCTGCTACACATTTAATCTCAAGTCACTT 580  
DB 158 CCTGAGAGCAGACCTTGTGAAGTCAAGAGCTGCTACACATTTAATCTCAAGTCACTT 217  
QY 581 TGAGATTGATGATTCAGAGAAAGAGATTCAGAGTTTGAAGCAAAATGCAGTGAACA 640  
DB 218 TGAGATTGATGATTCAGAGAAAGAGATTCAGAGTTTGAAGCAAAATGCAGTGAACA 277  
QY 641 GCAACCAATCCAACTGTGTAATTTGAAGACTTTGAGTCTTCACTTACCCGTTCCGTAACCC 700  
DB 278 GCAACCAATCCAACTGTGTAATTTGAAGACTTTGAGTCTTCACTTACCCGTTCCGTAACCC 337  
QY 701 CACTCTCAAGCCCTCTGAATATTCCTGCTTACCCCGAATATTTACAGTACATC 760  
DB 338 CACTCTCAAGCCCTCTGAATATTCCTGCTTACCCCGAATATTTACAGTACATC 397  
QY 761 TGCAGAGTCTCTTGGCCAG 781  
DB 398 TGCAGAGTCTCTTGGCCAG 418

RESULT 12  
US-10-029-386-15435  
Sequence 15435, Application US/10029386  
Publication No. US20030194704A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Hanzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
FILE REFERENCE: AEOMICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 34288  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 15435  
LENGTH: 379  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC021609.3  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.4  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.8  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.2  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.2  
OTHER INFORMATION: SWISSPROT HIT: Q9UPW0, EVALUE 1.80e+00  
OTHER INFORMATION: EST HUMAN HIT: AU132586.1, EVALUE 0.00e+00  
OTHER INFORMATION: NT HIT: g114729757, EVALUE 0.00e+00  
US-10-029-386-15435

Query Match 13.2%; Score 277; DB 16; Length 379;  
Best Local Similarity 99.5%; Pred. No. 1.9e-140;  
Matches 377; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 402 TTAAAGCTTGTGTGAGCCCTGAGATTCCTGAGCTTCGCGAGCCCTCAGAAAGCATTT 461  
DB 1 TTAAAGCTTGTGTGAGCCCTGAGATTCCTGAGCTTCGCGAGCCCTCAGAAAGCATTT 60  
QY 462 TAGGTCAAGCAGAGCAGAGAGATTAAGTGGCCCACTCCCGGTGGCATTCCTGATC 521  
DB 61 TAGGTCAAGCAGAGCAGAGAGATTAAGTGGCCCACTCCCGGTGGCATTCCTGATC 120  
QY 522 CTGAGAGCAGACCTTGTGAAGTCAAGAGCTGCTACACATTTAATCTCAAGTCACTT 581  
DB 121 CTCGAGAGCAGACCTTGTGAAGTCAAGAGCTGCTACACATTTAATCTCAAGTCACTT 180







```

: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.57
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.79
: OTHER INFORMATION: SWISSPROT HIT: O61608, EVALUAE 4.00e-04
: OTHER INFORMATION: EST HUMAN HIT: AA085543.1, EVALUAE 7.00e-94
: OTHER INFORMATION: NT HIT: g113325067, EVALUAE 5.00e-94
US-10-029-386-14338

```

Query Match	8.4%	Score 175	DB 16	Length 175
Best Local Similarity	100.0%	Pred. No.	1.7e-84	
Matches 175	Conservative 0	Mismatches 0	Indels 0	Gaps 0

QY	1767	AAAGAGTCGACGATTTCTCTTAAGCAGTGGAGTCCTAACTCATGATAAAGGTTCTCTC	1826
Db	175	AAAGAGTCGACGATTTCTCTTAAGCAGTGGAGTCCTAACTCATGATAAAGGTTCTCTC	116
QY	1827	AAGAGATGCTCTCTTGGGGGAGAGAAAGCCCGACAAAGTATGTACAAGACGACATCCA	1886
Db	115	AAGAGATGCTCTCTTGGGGGAGAGAAAGCCCGACAAAGTATGTACAAGACGACATCCA	56
QY	1887	GCTTCATGCGCAGGATGGCGAATCTCTCCAGGAGAAAGCGCATATTAT	1941
Db	55	GCTTCATGCGCAGGATGGCGAATCTCTCCAGGAGAAAGCGCATATTAT	1

Search completed: August 27, 2005, 17:33:32  
Job time : 902.113 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 23:18:31 ; Search time 4537.36 Seconds  
(without alignments)  
17558.328 Million cell updates/sec

Title: US-09-371-347A-47  
Perfect score: 2093  
Sequence: 1 atgaggaggttcgtctact.....ttcagatatttcgtcaca 2093

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues  
Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : EST:\*

1: gb\_est1: \*  
2: gb\_est2: \*  
3: gb\_hic: \*  
4: gb\_est3: \*  
5: gb\_est4: \*  
6: gb\_est5: \*  
7: gb\_est6: \*  
8: gb\_gsa1: \*  
9: gb\_gsa2: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1288	61.5	3100	3	BC062577 Homo sapi
2	956	45.7	3143	3	BC035977 Homo sapi
3	719	34.4	908	5	BX348674 BX348674
4	689	32.9	874	4	BM801462 AGENCOURT
5	623	29.8	646	7	CN260357 170004241
6	565	27.0	565	1	AU279788 AU279788
7	535	25.6	877	1	AU124440 AU124440
8	531	25.4	1061	5	BQ218755 AGENCOURT
9	517	24.7	826	4	BI772430 603055786
10	512	24.5	776	6	CB997527 AGENCOURT
11	507	24.2	834	5	BU941078 AGENCOURT
12	455	21.7	822	1	AU132586 AU132586
13	448	21.4	591	2	AW965709 EST377782
14	446	21.3	818	6	CD559384 AGENCOURT
15	434	20.7	591	4	BI025283 RC5-MT025
16	431	20.6	974	5	BX375211 BX375211
17	406	19.4	710	5	BU570323 AGENCOURT
18	384	18.3	527	7	BI025277 RC5-MT025
19	374	17.9	579	7	CN260360 170006001
20	368	17.6	852	5	BQ431497 AGENCOURT
21	361	17.2	692	7	CN260359 170004706
22	360	17.2	685	4	BM049352 603626120
23	359	17.2	499	6	CD704108 EST20635
24	354	16.9	386	1	AA279726 2892d10.r

25	341	16.3	395	4	BM838530 K-EST0114
26	340	16.2	526	2	AW952883 EST364953
27	337	16.1	818	7	CF995233 AGENCOURT
28	335	16.0	413	2	BF810368 BF810368
29	335	16.0	413	2	BF810479 RC5-C1014
30	335	16.0	413	2	AA085543 AA085543
31	331	15.8	521	6	CB164340 K-EST0225
32	320	15.3	839	4	BG531787 602560355
33	311	14.9	478	4	BM754488 K-EST0031
34	302	14.4	440	4	BG877205 OV3-HT046
35	297	14.2	416	6	CB996520 AGENCOURT
36	292	14.0	528	2	BE301292 BA89D07.X
37	291	13.9	664	7	CR768694 DKF2D459K
38	291	13.9	667	7	CR770923 DKF2D463N
39	291	13.9	667	7	CR557482 DKF2D469K
40	276	13.2	481	7	CR549172 DKF2D459J
41	272	13.0	301	1	AL704780 DKF2D686M
42	269	12.9	642	2	BF346446 602020302
43	264	12.6	366	6	CB298361 220019.re
44	257	12.3	366	2	BF808461 OV1-C1017
45	257	12.3	368	1	AA355001 EST63417

## ALIGNMENTS

RESULT 1	BC062577	3100 bp	mRNA	linear	HTC 25-NOV-2003
LOCUS	BC062577				
DEFINITION	Homo sapiens cDNA clone IMAGE:5189058, containing frame-shift errors.				
ACCESSION	BC062577				
VERSION	BC062577.1	GI:38511756			
KEYWORDS	HTC.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Strausberg, R.L., Peingold, R.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Schenker, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Scheffer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Datchenko, L., Mardina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uebelin, T.B., Toshitani, S., Carninci, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McMan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hilyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, U., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, V.S., Krzywinski, M.I., Skalska, U., Smallus, D.B., Scherch, A., Schein, J.B., Jones, S.J. and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
TITLE	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
JOURNAL	22388257				
MEDLINE	12477932				
PubMed	2 (bases 1 to 3100)				
REFERENCE	Strausberg, R.				
AUTHORS	Direct Submission				
TITLE	Submitted (24-NOV-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,				
JOURNAL	USA NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>				
REMARK	Contact: MGC help desk				
COMMENT	Email: <a href="mailto:cgabs-r@mail.nih.gov">cgabs-r@mail.nih.gov</a>				
	Tissue Procurement: Life Technologies, Inc.				

CDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland:  
 Web site: <http://www.nisc.nih.gov/>  
 Contact: nisc\_mgc@nhi.nih.gov  
 Ahlert, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
 Blakesley, R.M., Boufield, G.G., Breen, K., Brinkley, C., Brooks, S.,  
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,  
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,  
 Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McLoake, J.C.,  
 McDowell, J., Pearson, R., Stancik, S., Thomas, P.J., Touchman, J.W.,  
 Teuregon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
 Series: IRAK Plate: 135 Row: e Column: 21  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA g1: 4505278  
 This clone has the following problem: frame shifted.

## FEATURES

## source

1. 3100  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5189058"  
 /issue\_type="Colon, Kidney, Stomach, adult, whole pooled"  
 /clone\_lib="NIH MGC\_116"  
 /lab\_host="DH10B"  
 /note="Vector: PCMV-SPORT6"

## ORIGIN

Query Match 51.5%; Score 1288; DB 3; Length 3100;  
 Best Local Similarity 99.6%; Pred. No. 0;  
 Matches 1808; Conservative 0; Mismatches 2; Indels 5; Gaps 2;

283 GGTCTCGGTGATTCAGAAATACCTTCTTTCAGATGCGGAGAAATATGATTAACGA 342  
 172 GGTCTCGGTGATTCAGAAATACCTTCTTTCAGATGCGGAGAAATATGATTAACGA 231  
 343 CTTCAAGAGCTTGAAGCCCGCATTTCTATGACACTGACATGACATGACATGACATG 402  
 232 CTTCAAGAGCTTGAAGCCCGCATTTCTATGACACTGACATGACATGACATGACATG 291  
 403 TTGAACTTGTGTTGAGCGGTGATTTGCTGACATGCTGCGCAGCCCTCAGAAAGCATTT 462  
 292 TTGAACTTGTGTTGAGCGGTGATTTGCTGACATGCTGCGCAGCCCTCAGAAAGCATTT 351  
 463 AGGTCAAGCAGAGCAGAGAGAGATTAAGTGGCGCACTCCCGGTGACATCACTGCATCC 522  
 352 AGGTCAAGCAGAGCAGAGAGAGATTAAGTGGCGCACTCCCGGTGACATCACTGCATCC 411  
 523 TTGAGGACAGACTTGTGAAATGACAGAGTGTACATTTGAATTTCAAGTGCAGCTTCTG 582  
 412 TCGAGGACAGACTTGTGAAATGACAGAGTGTACATTTGAATTTCAAGTGCAGCTTCTG 471  
 583 AGATTCAGTATTCAGAGAAAGAGATTTCTGAGGTTTGAAGCAAAATGAGTGAAGC 642  
 472 AGATTCAGTATTCAGAGAAAGAGATTTCTGAGGTTTGAAGCAAAATGAGTGAAGC 531  
 643 AACCAATCCATGTTGTAATTTGAAGACTTTGAGTCTTCACTTACCCGTTGCTGATCCCA 702  
 532 AACCAATCCATGTTGTAATTTGAAGACTTTGAGTCTTCACTTACCCGTTGCTGATCCCA 591  
 703 CTCTCAAGAGCTTCTGAAATTTCTGTTTACCCCAAGATTTTTCAGAGTACATCTG 762  
 592 CTCTCAAGAGCTTCTGAAATTTCTGTTTACCCCAAGATTTTTCAGAGTACATCTG 651  
 763 CAGAGTCTCTTGGCAGAGAGAAAGCAAGTATCTGATCTTGAAGAGATCCAGTTT 822  
 652 CAGAGTCTCTTGGCAGAGAGAAAGCAAGTATCTGATCTTGAAGAGATCCAGTTT 711

823 CAGTGCATTTTCAAGGACGTTCACTTACGATGATGCGATTAACCACTCTG 882  
 712 CAGTGCATTTTCAAGGACGTTCACTTACGATGATGCGATTAACCACTCTG 771  
 883 CTGTGAGATTTGACATTTCAATACAGCTTTCTATACCTCTGAGATGCTTCTGAC 942  
 772 CTGTGAGATTTGACATTTCAATACAGCTTTCTATACCTCTGAGATGCTTCTGAC 831  
 943 GTGATCTGCTTACAGTATTTCTGAGTACAAAGCTTCTCTCAAGACTGAGTTGAA 1002  
 832 GTGATCTGCTTACAGTATTTCTGAGTACAAAGCTTCTCTCAAGACTGAGTTGAA 891  
 1003 GATTAAGAGAGCACTGCTCTTTTGAATTAAGGACAGACAAAGAAAGAGAGCT 1062  
 892 GATTAAGAGAGC-CTGCTCTTTTGAATTAAGGACAGACAAAGAAAGAGAGCT 950  
 1063 ACCCTTACCCAGATTTACCTGCGGAGATCTCTCCCACTTCTTCTGCTCTT 1122  
 951 ACCCTTACCCAGATTTACCTGCGGAGATTTCTCTCAAGTATTTTACCTGCTCTT 1010  
 1123 GAAATCCGACATTTCTTAAAGGATTTTGGAGCCCTTGTGACTATACAGTAC 1182  
 1011 GAAATCCGACATTTCTTAAAGGATTTTGGAGCCCTTGTGACTATACAGTAC 1070  
 1183 AGTCTGAAAGGAGGCTTACAGAGCTGTGCAATTAACAAAGGAGCCGATTTATAC 1242  
 1071 AGTCTGAAAGGAGGCTTACAGAGCTGTGCAATTAACAAAGGAGCCGATTTATAC 1130  
 1243 CGTTTGTACAGATGCTGCTGCTGCTTGTGATCTCTCTGCTTCTCTCTTCTG 1302  
 1131 CGTTTGTACAGATGCTGCTGCTGCTTGTGATCTCTCTGCTTCTCTCTTCTG 1190  
 1303 CAGCCACCACTAGTCTCTGCTGCAATCTTCTTAACTTCAACCCAGACATATTCG 1362  
 1191 CAGCCACCACTAGTCTCTGCTGCAATCTTCTTAACTTCAACCCAGACATATTCG 1250  
 1363 TGTGCAAGCTCAAGTTTATTTTCAACGAGAAAGCTCATTTGTCTTCAACATTTGGA 1422  
 1251 TGTGCAAGCTCAAGTTTATTTTCAACGAGAAAGCTCATTTGTCTTCAACATTTGGA 1310  
 1423 TTTCTGTACAGGCAACAAGAGGTTTGGGAAAGGAGTATGACAGGCTGCTG 1482  
 1311 TTTCTGTACAGGCAACAAGAGGTTTGGGAAAGGAGTATGACAGGCTGCTG 1370  
 1483 TTGTTGTTGCTTCACTTCTTCAAGCAATCATGATCCATGAGTGAAGCAGCGGAAA 1542  
 1371 TTGTTGTTGCTTCACTTCTTCAAGCAATCATGATCCATGAGTGAAGCAGCGGAAA 1430  
 1543 GCCCTGCTCTTAAGATATTCATCTCTCTGAAACAATTTCTTCACTTACAGAT 1602  
 1431 GCCCTGCTCTTAAGATATTCATCTCTCTGAAACAATTTCTTCACTTACAGAT 1490  
 1603 GAGCCCTCAATCCCACTCAATGATGAGTCCAGAAACCGGCAATACCCGTTATTTGG 1662  
 1491 GAGCCCTCAATCCCACTCAATGATGAGTCCAGAAACCGGCAATACCCGTTATTTGG 1550  
 1663 TTCTTCAACAT---AGAAATCTCAAGAAACAACCCAGATGAAATTTTGGAGCAATG 1718  
 1551 TTCTTCAACATGAGAGAAATCTCAAGAAACAACCCAGATGAAATTTTGGAGCAATG 1610  
 1719 TGGTTGTTTGGCTGAGCAATTAAGATGAGATTTATTTCAAGAAAGCTCAGA 1778  
 1611 TGGTTGTTTGGCTGAGCAATTAAGATGAGATTTATTTCAAGAAAGCTCAGA 1670  
 1779 CATTTCTTAAGCAATGATTTTCACTTAAAGTTTCTTCTCAAGATGCTCTT 1838  
 1671 CATTTCTTAAGCAATGATTTTCACTTAAAGTTTCTTCTCAAGATGCTCTT 1730  
 1839 GTTGGGAGAGAGAGCCCAAGCAAGATATGTAAGCAACAATCCAGTTCATGGCCAG 1898  
 1731 GTTGGGAGAGAGAGCCCAAGCAAGATATGTAAGCAACAATCCAGTTCATGGCCAG 1790

QY	1899	CAGGTGGCGAAGATCCCTCCACAGAGAAGGGCCATATTTATGTGTGGAGATGCAAG	1958
Db	1791	CAGGTGGCGAAGATCCCTCCACAGAGAAGGGCCATATTTATGTGTGGAGATGCAAG	1850
QY	1959	AATATGGCCGAAGATGTACATGATGCCCTGTGTGCAATTAATTAAGCAAGAGTTGAGTT	2018
Db	1851	AATATGGCCGAAGATGTACATGATGCCCTGTGTGCAATTAATTAAGCAAGAGTTGAGTT	1910
QY	2019	GAAGAACTAGAGCAATGAAAAACCTTGCCCACTTTAAAAAGAGAAAAACGTAACCTTCAG	2078
Db	1911	GAAGAACTAGAGCAATGAAAAACCTTGCCCACTTTAAAAAGAGAAAAACGTAACCTTCAG	1970
QY	2079	GATATTTGGTCATTA	2093
Db	1971	GATATTTGGTCATTA	1985
RESULT 2			
LOCUS	BC035977		
DEFINITION	Homo sapiens, clone IMAGE:4611253, mRNA.	3143 bp	linear
ACCESSION	BC035977		
VERSION	BC035977.1	GI:23243305	
KEYWORDS	HTC.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
REFERENCE	1 (bases 1 to 3143)		
AUTHORS	Strausberg,R.		
TITLE	Direct Submission		
JOURNAL	Submitted (31-JUL-2002)	National Institutes of Health, Mammalian	

REMARK	COMMENT
NIH-MGC Project URL: <a href="http://imgc.nci.nih.gov">http://imgc.nci.nih.gov</a>	
Contact: MGC help desk	
Email: <a href="mailto:gcgabs-remail.nih.gov">gcgabs-remail.nih.gov</a>	
Tissue Procurement: CLONTECH	
cDNA Library Preparation: CLONTECH Laboratories, Inc.	
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)	
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305	
Web site: <a href="http://www.shgc.stanford.edu">http://www.shgc.stanford.edu</a>	
Contact: (Dickson, Mark) <a href="mailto:mcd@paxill.stanford.edu">mcd@paxill.stanford.edu</a>	
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.	
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a>	
Series: IRAL Plate: 41 Row: 9 Column: 2	
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4505278	
This clone has the following problem: frame shifted.	
Location/Qualifiers	

FEATURES	SOURCE
Location/Qualifiers	1. .3143
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone_image="4611253"
	/tissue_type="Kidney"
	/clone_id="NH_MGC_75"
	/lab_host="DH10B"
	/note="Vector: pDNR-LIB"

```

Query Match Similarity      95.7%   Score 956; DB 3; Length 3143;
      Best Local Similarity  99.8%   Pred. No. 0;
Matches 1056; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy      1  ATGAGGAGGTTCTTGTTACTATATGCTACACGACGAGGACGACGACGACGATCGCAGAA 60
Ob      52  ATGAGGAGAGCTTCTGTTACTATATGCTACACGACGAGGACGACGACGACGATCGCAGAA 111

```

OY	6	GAATGTGTGACGAAGCTGTGTGACATGAGATTTTCTGACAGTCTTCACTGATTAATGAA	120
Db	112	GAATTAATGTAGCAAGCTGTGTGACATGAGATTTTCTGACAGTCTTCACTGATTAATGAA	171
OY	121	TCCGATAAGTATGACTTAAACCGAAGCAGCTCTTGTGTGTGTCTTACACG	180
Db	172	TCCGATAAGTATGACTTAAACCGAAGCAGCTCTCTGTGTGTGTGTCTTACACG	231
OY	181	GGCACCGGAGACCAACCCGACACAGCCCGCACTTTGTAAAGAAATACAGAACCAACA	240
Db	232	GGCACCGGAGACCAACCCGACACAGCCCGCAAGTTGTAAAGAAATACAGAACCAACA	291
OY	241	CTGCGGCTGATTTCTTGTGCTCACTGCGGTATGGTTACGTGGTCTCGGATTCAGAA	300
Db	292	CTGCGGCTGATTTCTTGTGCTCACTGCGGTATGGTTACGTGGTCTCGGATTCAGAA	351
OY	301	TACACCTACTTTTGCAATGGGGGGAGATTAATGATTAACGACTTCAAGACTTGAAGCC	360
Db	352	TACACCTACTTTTGCAATGGGGGGAGATTAATGATTAACGACTTCAAGACTTGAAGCC	411
OY	361	CGGATTTCTATGACAATGACACATGACAGATGACGTGTAGGTTTGAACCTTGTGTGTAG	420
Db	412	CGGATTTCTATGACACTGACACATGACAGATGACGTGTAGGTTTGAACCTTGTGTGTAG	471
OY	421	CCGTGATTTGCTGACCTCTGGCAGCCCTCGAAGACATTTTATAGTCAAGCAGACGACAA	480
Db	472	CCGTGATTTGCTGACCTCTGGCAGCCCTCGAAGACATTTTATAGTCAAGCAGACGACAA	531
OY	481	GAGGAGATTAAGTGGCGCACTCCCGGTGGCATCACTGCACTCTTGAGACAGACCTTGTG	540
Db	532	GAGGAGATTAAGTGGCGCACTCCCGGTGGCATCACTGCACTCTTGAGAGACAGACCTTGTG	591
OY	541	AAGTCAGAGCTGCTACCATTTGAATCTCAAGTGAGCTTGGAGATTCGATGATTCAGGA	600
Db	592	AAGTCAGAGCTGCTACCATTTGAATCTCAAGTGAGCTTGGAGATTCGATGATTCAGGA	651
OY	601	AGAAAGATTCCTGAGTTTGTGAAGCAAAATGACATGAAACAGACAAACCAATCCATGTGTGA	660
Db	652	AGAAAGATTCCTGAGTTTGTGAAGCAAAATGACATGAAACAGACAAACCAATCCATGTGTGA	711
OY	661	ATTGAAGACTTTGAGTCTCACTTACCCGTTGGTACCCCACTCTCAACAAGCCTCTCTG	720
Db	712	ATTGAAGACTTTGAGTCTCTCACTTACCCGTTGGTACCCCACTCTCAACAAGCCTCTCTG	771
OY	721	AATATTCCTGATTAAACCCCAAGATTTTACAGGTAATCTGACAGAGTCTCTTGGCCAG	780
Db	772	AATATTCCTGATTAAACCCCAAGATTTTACAGGTAATCTGACAGAGTCTCTTGGCCAG	831
OY	781	GAGAAAGCCCAAGATCTGTGACTTACAGCAGATCCAGTTTTCAGTGTCCAAATTTCAAAG	840
Db	832	GAGAAAGCCCAAGATCTGTGACTTACAGCAGATCCAGTTTTCAGTGTCCAAATTTCAAAG	891
OY	841	GCAATTGACCTTAACAGAAATGAGGCCATTAACCAACTGCTGCTGTAGAAATTTGGACATT	900
Db	892	GCAATTGACCTTAACAGAAATGAGGCCATTAACCAACTGCTGCTGTAGAAATTTGGACATT	951
OY	901	TCAATATCAGACTTTTCTTATCAGCTGAGATGCTTACAGCTGATCTGCTTAAACAGT	960
Db	952	TCAATATCAGACTTTTCTTATCAGCTGAGATGCTTACAGCTGATCTGCTTAAACAGT	1011
OY	961	GATTCTGAGTACAAAGCTTACTCCAAAGACTGACGCTTGAAGTAAAGAGAGACCTGC	1020
Db	1012	GATTCTGAGTACAAAGCTTACTCCAAAGACTGACGCTTGAAGTAAAGAGAGACCTGC	1071
OY	1021	GTCCCTTTGAATAAAGGACGACACAAAGAAAGG	1058
Db	1072	GTCCCTTTGAATAAAGGACGACACAAAGAAAGG	1109

RESULT 3	
BX348674	
LOCUS	
BX348674	908 bp
	mRNA
	linear
	EST 08-APR-2004

DEFINITION BX348674 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
CDNA clone CS0DC010Y11 5-PRIME, mRNA sequence.  
ACCESSION BX348674  
VERSION BX348674.1 GI:30375301  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
AUTHORS Li, W.-B., Gruber, C., Jesssee, J., and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
3392.f  
For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?c=CS0BAG006ZB02\_CS00490\_1&c=3392.f

FEATURES  
source  
Location/Qualifiers  
1..908  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DC010Y11"  
/issue\_type="NEUROBLASTOMA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN  
Query Match 34.4%; Score 719; DB 5; Length 908;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 719; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 674 AGTCTCACTTACCCGTTGGTACCACCTCTCAGAGCTCTCTGATATTTCTGGT 733  
DB 28 AGTCTCACTTACCCGTTGGTACCACCTCTCAGAGCTCTCTGATATTTCTGGT 87  
QY 734 TACCCCAAGATATTTACAGTACATCTGACAGAGTCTCTTGGCCAGAGAAAGCCAG 793  
DB 88 TACCCCAAGATATTTACAGTACATCTGACAGAGTCTCTTGGCCAGAGAAAGCCAG 147  
QY 794 TATCTGTGATTCAGAGATCCAGTTTTCAGAGCAATTTCAAGAGGAGTTCAACTTA 853  
DB 148 TATCTGTGATTCAGAGATCCAGTTTTCAGAGCAATTTCAAGAGGAGTTCAACTTA 207  
QY 854 CTACGAATGATGCAATTAACCACTCTGCTGTGATTAATGACATTTCAATACAGACT 913  
DB 208 CTACGAATGATGCAATTAACCACTCTGCTGTGATTAATGACATTTCAATACAGACT 267  
QY 914 TTTTCTATCAGCTGAGATGCTTTCAGAGTATCTCCCTTAACAGTATTCAGAGTAC 973  
DB 268 TTTTCTATCAGCTGAGATGCTTTCAGAGTATCTCCCTTAACAGTATTCAGAGTAC 327  
QY 974 AAAGCTACTCCAAAGACTGAGCTTAAGATTAAGAGAGAGAGCTGCTCTTTGAAA 1033  
DB 328 AAAGCTACTCCAAAGACTGAGCTTAAGATTAAGAGAGAGAGCTGCTCTTTGAAA 387  
QY 1034 TAAAGGAGACACAAAGAGAGAGAGCTTACCCAGCATATATCTGCGGAGATGT 1093  
DB 388 TAAAGGAGACACAAAGAGAGAGAGCTTACCCAGCATATATCTGCGGAGATGT 447  
QY 1094 CTCTCCAGTTCAATTTTAACTGCTGTCTTGAATCCAGAGCAATTTCTAAAGAGCATTTT 1153

DB 448 CTCTCCAGTTCAATTTTAACTGCTGTCTTGAATCCAGAGCAATTTCTAAAGAGCATTTT 507  
QY 1154 TCGAGACCCCTTGTGACTATACAGTACAGAGTCTGTAAGAGGACAGCTACAGAGACTGT 1213  
DB 508 TCGAGACCCCTTGTGACTATACAGTACAGAGTCTGTAAGAGGACAGCTACAGAGACTGT 567  
QY 1214 GCAGTAAACAGAGGAGAGCCGATTAATGCGCTTTGTACAGATGCTGCTGCTTGT 1273  
DB 568 GCAGTAAACAGAGGAGAGCCGATTAATGCGCTTTGTACAGATGCTGCTGCTTGT 627  
QY 1274 TGGATCTCTCTCGCTTCCCTTCTTGCAGACACCATCTCCTGCTGGAATTC 1333  
DB 628 TGGATCTCTCTCGCTTCCCTTCTTGCAGACACCATCTCCTGCTGGAATTC 687  
QY 1334 TTCTTAACCTTCAACCCAGACCATATATTCGTGCAAGCTCAAGTTATTTCAACCCAGA 1392  
DB 688 TTCTTAACCTTCAACCCAGACCATATATTCGTGCAAGCTCAAGTTATTTCAACCCAGA 746

RESULT 4  
BM801462 874 bp mRNA 11near EST 05-MAR-2002  
LOCUS  
DEFINITION AGENCOURT\_6459212 NIH\_MGC\_88 Homo sapiens cDNA clone IMAGE:5560477  
5', mRNA sequence.  
ACCESSION BM801462  
VERSION BM801462.1 GI:19118285  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLNL12286 row: 1 column: 14  
High quality sequence stop: 710.

FEATURES  
source  
Location/Qualifiers  
1..874  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5560477"  
/issue\_type="duodenal adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_88"  
/note="Organ: small intestine; Vector: pCMV-SPORT6;  
Site 1: NotI; Site 2: SalI; Cloned unidirectionally;  
oligo-dT primed. Average insert size 1.767 kb. Library  
enriched for full-length clones and constructed by Life  
Technologies. Note: this is a NIH\_MGC Library."

ORIGIN  
Query Match 32.9%; Score 689; DB 4; Length 874;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 729; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGAGGTTTCTGTACTATATGCTTACACAGAGGAGACGCAAGGCAATCGCAGAA 60  
DB 50 ATGAGAGGTTTCTGTACTATATGCTTACACAGAGGAGACGCAAGGCAATCGCAGAA 109  
QY 61 GAAATGTGAGCAACCTGTGTACATGATTTTTCGAGATTCATGATATTAAGGAA 120

Db 110 GAAATATGTGACCAAGCTGTGTGATCATGATTTCTGCAGATCTTCACTGATTTAGTGA 169  
Qy 121 TCCGATTAAGTATGACCTAAACCCGAAACAGCTCCCTGTTGTTGTGTTCTACAG 180  
Db 170 TCCGATTAAGTATGACCTAAACCCGAAACAGCTCCCTGTTGTTGTGTTCTACAG 229  
Qy 181 GGCACCGGAGACCCACCCGACACAGCCCGCAAGTTGTTAGAAATACAGAACCAACA 240  
Db 230 GGCACCGGAGACCCACCCGACACAGCCCGCAAGTTGTTAGAAATACAGAACCAACA 289  
Qy 241 CTGCGGTTGATTTCTTGTCTCACTGCGGTATGAGTTACTGGTCTCGGTATTCAGA 300  
Db 290 CTGCGGTTGATTTCTTGTCTCACTGCGGTATGAGTTACTGGTCTCGGTATTCAGA 349  
Qy 301 TACACCTACTTTTGCAATGGGGGAAAGATTAATTGAATTAACGACTTCAGAGCTTGAGCC 360  
Db 350 TACACCTACTTTTGCAATGGGGGAAAGATTAATTGAATTAACGACTTCAGAGCTTGAGCC 409  
Qy 361 CGGCAATTTCTATGACACTGACATGACATGATGATGATGATGATGATGATGATGATGAT 420  
Db 410 CGGCAATTTCTATGACACTGACATGACATGATGATGATGATGATGATGATGATGATGAT 469  
Qy 421 CGGTGATTTGCTGACCTGCGCAAGCTTCAGAAACATTTTATGATGATGATGATGATGAT 480  
Db 470 CGGTGATTTGCTGACCTGCGCAAGCTTCAGAAACATTTTATGATGATGATGATGATGAT 529  
Qy 481 GAGGAGATTAAGTGGGCACTCCCGGTGATGATGATGATGATGATGATGATGATGATGAT 540  
Db 530 GAGGAGATTAAGTGGGCACTCCCGGTGATGATGATGATGATGATGATGATGATGATGAT 589  
Qy 541 AAGTCAGAGCTGCTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
Db 590 AAGTCAGAGCTGCTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 649  
Qy 601 AGAAAGATTTGAGGTTTGAAGCAAAATGACATGACAGCAACCAATCCATGTTGTA 660  
Db 650 AGAAAGATTTGAGGTTTGAAGCAAAATGACATGACAGCAACCAATCCATGTTGTA 709  
Qy 661 ATTGAAGACTTTGAGTCTTCACTTACCCGTTGGTACCCCACTGTCAGAACCTTCTG 720  
Db 710 ATTGAAGACTTTGAGTCTTCACTTACCCGTTGGTACCCCACTGTCAGAACCTTCTG 769  
Qy 721 AATATTCCTGTTTACCCCT 740  
Db 770 AATATTCCTGTTTACCCCT 789

RESULT 5  
CN260357 646 bp mRNA linear EST 16-MAY-2004  
LOCUS 17000424179730 GRN\_BS Homo sapiens cDNA 5', mRNA sequence.  
DEFINITION CN260357  
ACCESSION CN260357.1 GI:47276771  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Butelostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 646)  
Brandenberger, R., Wei, H., Zhang, S., Lei, S., Muraue, J., Flak, G.J.,  
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,  
Lebkowski, J. and Stanton, L.W.  
Transcriptome characterization elucidates signaling networks that  
control human ES cell growth and differentiation  
Nat. Biotechnol. 22 (6), 707-716 (2004)  
Contact: Brandenberger R  
Regenerative Medicine  
Geron Corporation  
230 Constitution Drive, Menlo Park, CA 94025, USA  
Tel: 650 473 8658  
Fax: 650 473 7760  
Email: rbrandenberger@geron.com  
Insert Length: 646 Std Error: 0.00.

FEATURES  
source

Location/Qualifiers  
1..646  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="embryonic stem cells, cell lines H1, H7, and H9"  
/clone\_lib="GRN BS"  
/note="Oligo dt primed, full-length enriched cDNA library from undifferentiated hbs cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions"

## ORIGIN

Query Match 29.8%; Score 623; DB 7; Length 646;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 623; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 987 AAGCTGACCTTGAAGATTAAGAGAGACACTGCTCTTTGAAATTAAGGACAGAC 1046  
Db 24 AAGCTGACCTTGAAGATTAAGAGAGACACTGCTCTTTGAAATTAAGGACAGAC 83  
Qy 1047 AAAGAAAGAGAGCTTACCTTACCCAGATATACCTGGGATGTTCTCTCACTTAT 1106  
Db 84 AAAGAAAGAGAGCTTACCTTACCCAGATATACCTGGGATGTTCTCTCACTTAT 143  
Qy 1107 TTTTACCTGCTGCTTGAATCCAGCAATTCCTAATAAGGCAATTTTGCGAGCCCTGT 1166  
Db 144 TTTTACCTGCTGCTTGAATCCAGCAATTCCTAATAAGGCAATTTTGCGAGCCCTGT 203  
Qy 1167 GGAATTAACAGTGAAGTGTGTAAGAGGCAAGGCTGACAGAGCTGTGACATTAACAG 1226  
Db 204 GGAATTAACAGTGAAGTGTGTAAGAGGCAAGGCTGACAGAGCTGTGACATTAACAG 263  
Qy 1227 GGCAGCGATTAAGCGCTTTGACAGATGCTGTGCTGCTGTTGATCTCTCT 1286  
Db 264 GGCAGCGATTAAGCGCTTTGACAGATGCTGTGCTGCTGTTGATCTCTCTCT 323  
Qy 1287 GCGTTTCCCTTCTGCGAGCCAGCACTGATCTCCGCTCGAACAATCTTCAAACTTCA 1346  
Db 324 GCGTTTCCCTTCTGCGAGCCAGCACTGATCTCCGCTCGAACAATCTTCAAACTTCA 383  
Qy 1347 ACCAGACCATTAATGCTGACAGCTCAAGTTATTTCAACCGAGAAAGCTCATTTGT 1406  
Db 384 ACCAGACCATTAATGCTGACAGCTCAAGTTATTTCAACCGAGAAAGCTCATTTGT 443  
Qy 1407 CTTCAACATTTGTAATTTCTGTCTTACTGCAACAAGAGTTCTGCGAGAGGATATG 1466  
Db 444 CTTCAACATTTGTAATTTCTGTCTTACTGCAACAAGAGTTCTGCGAGAGGATATG 503  
Qy 1467 TACAGGCTGGCTGGCTGTTGTTGCTTCAAGTTCTTCAAGCCAAATACATGATCCCA 1526  
Db 504 TACAGGCTGGCTGGCTGTTGTTGCTTCAAGTTCTTCAAGCCAAATACATGATCCCA 563  
Qy 1527 TGAAGACAGCGGAGAGCCCTGCTCTTAAGATTCATCTCTCTGAGAAACAATTC 1586  
Db 564 TGAAGACAGCGGAGAGCCCTGCTCTTAAGATTCATCTCTCTGAGAAACAATTC 623  
Qy 1587 TTTCACTTACCAATGACCCCT 1609  
Db 624 TTTCACTTACCAATGACCCCT 646

RESULT 6  
A0279788 565 bp mRNA linear EST 31-JUL-2003  
LOCUS A0279788 CHONS2 Homo sapiens cDNA clone CHONS2001448 5', mRNA  
DEFINITION A0279788  
ACCESSION A0279788  
VERSION A0279788.1 GI:28299015  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Butelostomi;



REFERENCE 1 (bases 1 to 565)  
AUTHORS Imabayashi, H., Mori, T., Gojo, S., Kiyono, T., Sugiyama, T., Irie, R.,  
Isogai, T., Hata, J., Tomoya, Y., and Umezawa, A.  
TITLE Redifferentiation of dedifferentiated chondrocytes and  
chondrogenesis of human bone marrow stromal cells via chondrosphere  
formation with expression profiling by large-scale cDNA analysis  
JOURNAL Exp. Cell Res. 288 (1), 35-50 (2003)  
MEDLINE 22760698  
PUBMED 12878157  
COMMENT Contact: Takao Isogai  
Genomics Laboratory  
Helix Research Institute  
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: genomics@hri.co.jp  
HRI human cDNA project, Sugiyama, T.; Makatsugu, A.; Irie, R.;  
Umezawa, A.; Fukuma, M.; Kusakari, S.; Hata, J.; Ishii, S.; Yamamoto, J.;  
Isogo, Y.; Saito, K.; Nakamura, Y.; Masuko, Y.; Nagai, K.; Isogai, T.  
HRI human cDNA project; cDNA library construction & 5'-end one  
pass sequencing: Helix Research Institute.  
location/Qualifiers  
FEATURES  
source 1..565  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CHONS2001448"  
/cell\_type="chondrocytes"  
/clone\_1lb="CHONS2"  
/note="Vector: pME18SFL3"

ORIGIN  
Query Match 27.0%; Score 565; DB 1; Length 565;  
Best Local Similarity 100.0%; Pred. No. 3.7e-299;  
Matches 565; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

819 TTTTCAAGTGCATTTCAAGGCACTTCACTACTACGATGATGCTTAATAAACAC 878  
1 TTTTCAAGTGCATTTCAAGGCACTTCACTACTACGATGATGCTTAATAAACAC 60  
879 TCGGCTGATGATTTGACATTTTCAATACAGACTTTTCTATCAGCTGAGATGCTT 938  
61 TCGGCTGATGATTTGACATTTTCAATACAGACTTTTCTATCAGCTGAGATGCTT 120  
939 CAGCGTATGCTGCTTCAAGATTTCTGAGTACAAAGCTTCTCAAGATGCTGAGCT 998  
121 CAGCGTATGCTGCTTCAAGATTTCTGAGTACAAAGCTTCTCAAGATGCTGAGCT 180  
999 TGAAGATTAAG 1058  
181 TGAAGATTAAG 240  
1059 AGCTACTTATCCCAAGATATATCTGGGAGATGTTCTCTCAGTTATTTTACCTG 1118  
241 AGCTACTTATCCCAAGATATATCTGGGAGATGTTCTCTCAGTTATTTTACCTG 300  
1119 TCTTGAATCCAGAGATTTCTTAATAAGGCAATTTTTCGAGCCCTTGTGACTATAC 1178  
301 TCTTGAATCCAGAGATTTCTTAATAAGGCAATTTTTCGAGCCCTTGTGACTATAC 360  
1179 TGAAGATGCTGAAGAGGAGGCTTACAGAGAGCTGTGAGTAAACAGAGGAGG 1238  
361 TGAAGATGCTGAAGAGGAGGCTTACAGAGAGCTGTGAGTAAACAGAGGAGG 420  
1239 TAGCGCTTTGTAGAGATGCTGTGCTGCTGTGTGAGATCTCTCTCCCTTTCCCTTC 1298  
421 TAGCGCTTTGTAGAGATGCTGTGCTGCTGTGTGAGATCTCTCTCCCTTTCCCTTC 480  
1299 TTGGCAGCAGCACTGAGTCTGCTGCTGCAACATCTTCTTAACCTTCAACCCAGACATA 1358  
481 TTGGCAGCAGCACTGAGTCTGCTGCTGCAACATCTTCTTAACCTTCAACCCAGACATA 540

ORIGIN  
Query Match 25.6%; Score 535; DB 1; Length 877;  
Best Local Similarity 100.0%; Pred. No. 1.4e-282;  
Matches 535; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1143 AAGGATATTTTTCGAGCCCTTGTGAGCTATACAGTACAGTGCAGAAAGGCGAGGCT 1202  
192 AAGGATATTTTTCGAGCCCTTGTGAGCTATACAGTACAGTGCAGAAAGGCGAGGCT 251  
1203 ACAGAGAGCTGTGAGTAAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1262  
252 ACAGAGAGCTGTGAGTAAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 311  
1263 TGGCTGCTTTGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1322  
312 TGGCTGCTTTGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 371  
1323 GCTCGAAGATCTTCTCAACCTTCAACCCAGAGAGAGAGAGAGAGAGAGAGAG 1382  
372 GCTCGAAGATCTTCTCAACCTTCAACCCAGAGAGAGAGAGAGAGAGAGAGAG 431  
1383 TCACCGAG 1442  
432 TCACCGAG 491

QY 1443 AGAGGTTCTCGGAAAGGAGTATGTACAGGCTGCGCTTGTGTTGCTTCACTTCT 1502  
 DB 492 AAGAGTTCTCGGAAAGGAGTATGTACAGGCTGCGCTTGTGTTGCTTCACTTCT 551  
 QY 1503 TCAGCCAAATCATCATGATCCCATGAAGACAGCGGAAAGCCCTGCTCTTAAGATATC 1562  
 DB 552 TCAGCCAAATCATCATGATCCCATGAAGACAGCGGAAAGCCCTGCTCTTAAGATATC 611  
 QY 1563 CATCTCTCTCGAACAACAAATCTTCCACTTACCAAGATAGACCCCTCAATCCCATCAT 1622  
 DB 612 CATCTCTCTCGAACAACAAATCTTCCACTTACCAAGATAGACCCCTCAATCCCATCAT 671  
 QY 1623 AATGGTGGTCCAGAACCGGCATAGCCCGTTATTGAGTCTCTCAACATAGA 1677  
 DB 672 AATGGTGGTCCAGAACCGGCATAGCCCGTTATTGAGTCTCTCAACATAGA 726  
 RESULT 8  
 LOCUS BQ218755 1061 bp mRNA linear EST 02-MAY-2002  
 DEFINITION AGENCOURT\_7565843 NIH\_MGC\_92 Homo sapiens cDNA clone IMAGE:6041670  
 5', mRNA sequence.  
 ACCESSION BQ218755  
 VERSION BQ218755.1 GI:20400155  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 1061)  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-r@mail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 plate: LLM13279 row: n column: 07  
 High quality sequence stop: 518.  
 Location/Qualifiers  
 1..1061  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6041670"  
 /tissue\_type="embryonal carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_1lb="NIH\_MGC\_92"  
 /note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
 Average insert size 2.5 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC Library."  
 ORIGIN  
 Query Match 25.4%; Score 531; DB 5; Length 1061;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-280; Indels 0; Gaps 0;  
 Matches 531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 995 AGCTTGAAGTAAAGAGACACCTGCTCTTGTGAATAAGGACAGACAAAGAGA 1054  
 DB 1 AGCTTGAAGTAAAGAGACACCTGCTCTTGTGAATAAGGACAGACAAAGAGA 60  
 QY 1055 AAGAGACTACCTTACCCACAGATATACCTGGGGATGTTCTCTCACTTCACTTTTACT 1114  
 DB 61 AAGAGACTACCTTACCCACAGATATACCTGGGGATGTTCTCTCACTTCACTTTTACT 120  
 QY 1115 GGATGCTGAAGATCCGAGCAATTCCTAAAGGATTTTGGAGCCCTTGTGACTATA 1174

DB 121 GGATGCTGAAGATCCGAGCAATTCCTAAAGGATTTTGGAGCCCTTGTGACTATA 180  
 QY 1175 CCAATGACAGTGTCTGTAAGAAAGGACAGGCTACAGAGCTGTGACTAAACAAGGGGACGGC 1234  
 DB 181 CCAATGACAGTGTCTGTAAGAAAGGACAGGCTACAGAGCTGTGACTAAACAAGGGGACGGC 240  
 QY 1235 ATTATAGCGCTTGTGAGAGATGCGTGTGCTGTGTGATCTCTCTGCTTCC 1294  
 DB 241 ATTATAGCGCTTGTGAGAGATGCGTGTGCTGTGTGATCTCTCTGCTTCC 300  
 QY 1295 CTCTTGGCAGCCACACACTCACTCTCTGCTGAAATCTTCTTAACCCAGAC 1354  
 DB 301 CTCTTGGCAGCCACACACTCACTCTCTGCTGAAATCTTCTTAACCCAGAC 360  
 QY 1355 CATATTCGTGTGACAGCTCAATTTATTTACCCAGGAAGCTCCATTTGCTTCAAC 1414  
 DB 361 CATATTCGTGTGACAGCTCAATTTATTTACCCAGGAAGCTCCATTTGCTTCAAC 420  
 QY 1415 TTGTGGAATTTCTGTCTACGCAACAAGAGTTCTGCGGAAGGAGTATGACAGCT 1474  
 DB 421 TTGTGGAATTTCTGTCTACGCAACAAGAGTTCTGCGGAAGGAGTATGACAGCT 480  
 QY 1475 GGCTGGCCTTGTGTGCTTCACTTCAAGCCAAACATACATGATCCC 1525  
 DB 481 GGCTGGCCTTGTGTGCTTCACTTCAAGCCAAACATACATGATCCC 531  
 RESULT 9  
 LOCUS B1772430 826 bp mRNA linear EST 25-SEP-2001  
 DEFINITION 603055786F1 NIH\_MGC\_122 Homo sapiens cDNA clone IMAGE:5205285 5',  
 mRNA sequence.  
 ACCESSION B1772430  
 VERSION B1772430.1 GI:15764008  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 826)  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-r@mail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 plate: LLM11514 row: 1 column: 22  
 High quality sequence stop: 824.  
 Location/Qualifiers  
 1..826  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5205285"  
 /lab\_host="DH10B"  
 /clone\_1lb="NIH\_MGC\_122"  
 /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;  
 Site 1: NotI; Site 2: BcoRI (destroyed); RNA source  
 anonymous pool of 24 week female lung, 16 week female  
 spleen, and 20-22 week male spleens. Library is oligo-dT  
 primed and directionally cloned (BcoRI site is destroyed  
 upon cloning). Average insert size 1.4 kb, insert size  
 range 1-3 kb. Library is normalized and enriched for  
 full-length clones and was constructed by C. Gruber  
 (Invitrogen). Research Genetics tracking code 026. Note:  
 this is a NIH\_MGC Library."  
 FEATURES  
 Source



```

Db          688  AGAAGATCTGA  701

RESULT  11
LOCUS   BU941078
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

      834 bp      mRNA      linear      EST 18-OCT-2002
AGENCOURT 10540067 NIH_MGC 128 Homo sapiens cDNA clone
IMAGE:6712893 5', mRNA_sequence.
BU941078
BU941078.1 GI:24129897
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 834)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: NCI
CDNA Library Preparation: Michael Brownstein Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LIML at:
http://image.llnl.gov
Plate: L10CM3022 row: e column: 21
High quality sequence stop: 586.
Location/Qualifiers
1..834

```

ORIGIN	Query Match	24.2%	Score 507;	DB 5;	Length 834;
	Best Local Similarity	99.7%;	Pred. No. 3.8e-267;		
	Matches 607;	Conservative	2;	Indels 0;	Gaps 0;
QY	379	GGACATGAGATGACTGCTGATGTTTAACCTTGCTGTGAGCCGTGATTCGTGGACCTC			438
Db	3	GGACATGAGATGACTGCTGCTGATTAACCTTGCTGTGAGCCGTGATTCGTGGACCTC			62
QY	439	TGGCCAGCCCTCAGAAAGCATTTTAAAGTCAGACAGAGACAAGAGATTAAGTGGCCCA			498
Db	63	TGGCCAGCCCTCAGAAAGCATTTTAAAGTCAGACAGAGACAAGAGATTAAGTGGCCCA			122
QY	499	CTCCCGGTGGCATCACTCTGCATCCTTGAAGACAGACCCTTGTGAAGTCAGAGCTGCTAAC			558
Db	123	CTCCCGGTGGCATCACTCTGCATCCTTGAAGACAGACCCTTGTGAAGTCAGAGCTGCTAAC			182

Oy	ATTGAATCTGAAGTCAGAGCTTCTGAGATTGGAATTGATTCAGGAAGAAAGATTCTGAGGTT	618
Db	ATTGAATCTGAAGTCAGAGCTTCTGAGATTGGAATTGATTCAGGAAGAAAGATTCTGAGGTT	242
Oy	TTGAGCAAAATGCAGTGAAACAGCAACCAATCCAAATGTGTAATTGAAAGACTTGTAGTCC	678
Db	TTGAGCAAAATGCAGTGAAACAGCAACCAATCCAAATGTGTAATTGAAAGACTTGTAGTCC	302
Oy	TCACTTACCCGTTGGGTACCCTCACTCAACACCTCTCGAATATTTCCGTGTTTACC	738
Db	TCACTTACCCGTTGGGTACCCTCACTCAACACCTCTCGAATATTTCCGTGTTTACC	362
Oy	CCAGAAATTTTACAGGTACATCTGCAAGAGTCTCTTGAGCAGAGAAAGCCAAAGTATCT	798
Db	CCAGAAATTTTACAGGTACATCTGCAAGAGTCTCTTGAGCAGAGAAAGCCAAAGTATCT	422
Oy	GTGACTTCAGCAGATCCAGTCTTTCAAGTCCAAATTTCAAAGGCAATTTCAACTTACTACG	858
Db	GTGACTTCAGCAGATCCAGTCTTTCAAGTCCAAATTTCAAAGGCAATTTCAACTTACTACG	482
Oy	AATAGATGCATTAATAAACCACTCTGCTGTGTAATTGACATTTTCAAAATACAGACTTTTCC	918
Db	AATAGATGCATTAATAAACCACTCTGCTGTGTAATTGACATTTTCAAAATACAGACTTTTCC	542
Oy	TATCAGCCTGGAGATGCTTCAAGCGTATCTGCCCTTAACAGTGAATTCGAGGTACAAGC	978
Db	TATCAGCCTGGAGATGCTTCAAGCGTATCTGCCCTTAACAGTGAATTCGAGGTACAAGC	602
Oy	CTACTCCAA 987	
Db	CTACTCCAA 611	

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
1 (bases 1 to 822)
Oca,T., Sugiyama,T., Ishii,S., Suzuki,Y., Salto,K., Yamamoto,J.,
Nishikawa,T., Nakamura,Y., Negai,T., Sugano,S., Masuno,Y. and
Isogai,T.
HRI human cDNA project (Oca,T., Sugiyama,T., Ishii,S., Suzuki,Y.,
Salto,K., Yamamoto,J., Nishikawa,T., Nakamura,Y., Negai,T.,
Sugano,S., Masuno,Y., Isogai,T.)
Unpublished (2000)
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-1975
Fax: 81-438-52-3986
Email: genomcshri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
Location/Qualifiers
1..822
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cclone="NT2RP4000141"
/cell_type="teratocarcinoma"
/cell_line="NT2"

```

## ORIGIN

/clone\_11b="NT2RP4"  
/note="Vector: pME18SFLJ, mRNA from NT2 neuronal precursor  
cells after 2-weeks retinoic acid (RA) induction"

Query Match 21.7%; Score 455; DB 1; Length 822;

Best Local Similarity 99.6%; Pred. No. 1.8e-238; Indels 0; Gaps 0;

Matches 555; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 283 GGTCTCGGATTCAGAAATACCTACTTTGGCAATGGGGGAAAGATTAATTGATTAACGA 342  
DB 182 GGTCTCGGATTCAGAAATACCTACTTTGGCAATGGGGGAAAGATTAATTGATTAACGA 241  
QY 343 CTTCAGAGCTTGAGAGCCCGGCAATTTCTATGACATGACATGCAATGATCTGTAGGT 402  
DB 242 CTTCAGAGCTTGAGAGCCCGGCAATTTCTATGACATGACATGCAATGATCTGTAGGT 301  
QY 403 TTGAATCTGTGTGAGCCGCGGATTTGCTGAGCTCTGGCCAGCCCTCAAGAAAGCAATTT 462  
DB 302 TTGAATCTGTGTGAGCCGCGGATTTGCTGAGCTCTGGCCAGCCCTCAAGAAAGCAATTT 361  
QY 463 AGGTCAAGCAGAGAGCAAGAGATTAAGTGGGCACTCCGGTGGCATCAGCTGATCC 522  
DB 362 AGGTCAAGCAGAGAGCAAGAGATTAAGTGGGCACTCCGGTGGCATCAGCTGATCC 421  
QY 523 TTGAGACAGACCTTTGTGAAGTCAAGAGCTGCTACATTTGAATCTCAAGTGGCTTTG 582  
DB 422 TCAGAGCAGACCTTTGTGAAGTCAAGAGCTGCTACATTTGAATCTCAAGTGGCTTTG 481  
QY 583 AGATTCGATGATTCAAGAAAGAGATTTGAGGTTTGAAGCAAAATGCAAGGACAGC 642  
DB 482 AGATTCGATGATTCAAGAAAGAGATTTGAGGTTTGAAGCAAAATGCAAGGACAGC 541  
QY 643 AACCAATCCAAATGTTGTAATGAAGCTTGAAGTCTCACTTACCCGTTGGTACCCCA 702  
DB 542 AACCAATCCAAATGTTGTAATGAAGCTTGAAGTCTCACTTACCCGTTGGTACCCCA 601  
QY 703 CTCTCAGACGCTCTCTGAATATTTCTGTGTTTACCCCAAGATTTTACAGGTACATCTG 762  
DB 602 CTCTCAGACGCTCTCTGAATATTTCTGTGTTTACCCCAAGATTTTACAGGTACATCTG 661  
QY 763 CAGAGCTCTTGGCCAGAGAAAGCAAGTATCTGTGACTTCAAGCATCCAGTTT 822  
DB 662 CAGAGCTCTTGGCCAGAGAAAGCAAGTATCTGTGACTTCAAGCATCCAGTTT 721  
QY 823 CAAATGCCAATTTCAA 839  
DB 722 CAAATGCCAATTTCAA 738

RESULT 13  
AM965709 591 bp mRNA linear EST 01-JUN-2000  
LOCUS AM965709  
DEFINITION MAGE resequences, MAGI Homo sapiens cDNA, mRNA sequence.  
ACCESSION AM965709  
VERSION AM965709.1 GI:8155545  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 591)  
Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaepard, R., Gay, C.,  
Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and  
Quackenbush, J.  
Assessment of gene expression patterns in a model of colon tumor  
metastasis using a 19,200 element cDNA microarray  
Unpublished (2000)  
JOURNAL COMMENT  
Contact: John Quackenbush  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 3528  
Fax: 301 838 0208

Email: johng@tigr.org  
Plate: 218  
Seq primer: Reverse.  
Location/Qualifiers  
1..591

## FEATURES

source

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_11b="MAGE resequences, MAGI"  
/note="Vector: pBluescriptSKm"

## ORIGIN

Query Match 21.4%; Score 448; DB 2; Length 591;

Best Local Similarity 100.0%; Pred. No. 1.3e-234; Indels 0; Gaps 0;

Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1192 AAGCGAGGCTACAGAGCTGTCAGTAAACAAGGGGACGCCGATTAATAGCCGCTTGT 1251  
DB 1 AAGCGAGGCTACAGAGCTGTCAGTAAACAAGGGGACGCCGATTAATAGCCGCTTGT 60  
QY 1252 CGAGATGCTGTGCTGCTGCTTGTGATCTCTCTGCTTCCCTTCTTGCCAGCCCA 1311  
DB 61 CGAGATGCTGTGCTGCTGCTTGTGATCTCTCTGCTTCCCTTCTTGCCAGCCCA 120  
QY 1312 CTGAGTCTCTGCTGCAACATCTTCTTAACTTCAACCCAGACATATTCGTGCAAGC 1371  
DB 121 CTGAGTCTCTGCTGCAACATCTTCTTAACTTCAACCCAGACATATTCGTGCAAGC 180  
QY 1372 TCAGATTTATTCACCCAGAAAGCTCAATTTGTCTTCAACATGTGCAATTTCTGTCT 1431  
DB 181 TCAGATTTATTCACCCAGAAAGCTCAATTTGTCTTCAACATGTGCAATTTCTGTCT 240  
QY 1432 ACTGCCAACAAGAGGTTCTGCGAAGGAGATGTAACAGCTGAGCTGCTTGTGTT 1491  
DB 241 ACTGCCAACAAGAGGTTCTGCGAAGGAGATGTAACAGCTGAGCTGCTTGTGTT 300  
QY 1492 GCTTCAGTCTTTCAGCAACATCATGATCCCATGAAAGACGCGGAAAGCCCTGGCT 1551  
DB 301 GCTTCAGTCTTTCAGCAACATCATGATCCCATGAAAGACGCGGAAAGCCCTGGCT 360  
QY 1552 CCTAGATATTCATCTCTCTGCAACACAAATTTCTTCACTTACAGATGACCCCTCA 1611  
DB 361 CCTAGATATTCATCTCTCTGCAACACAAATTTCTTCACTTACAGATGACCCCTCA 420  
QY 1612 ATCCCATCATTAATGTTGGGTCAGGAA 1639  
DB 421 ATCCCATCATTAATGTTGGGTCAGGAA 448

RESULT 14  
CD559384 818 bp mRNA linear EST 11-JUN-2003  
LOCUS CD559384  
DEFINITION AGENCOURT 14401607 NIH-MGC 191 Homo sapiens cDNA clone  
IMAGE:30409775 5', mRNA sequence.  
ACCESSION CD559384  
VERSION CD559384.1 GI:31585452  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 818)  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
JOURNAL COMMENT  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics / NIH  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cgaabs-r@mail.nih.gov  
Tissue Procurement: Narayan Bhat  
cDNA Library Preparation: Clontech Laboratories, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/MLN at:  
<http://image.jnl.gov>  
 Plate: NDCM198 row: n column: 24  
 High quality sequence stop: 484.  
 Location/Qualifiers

## FEATURES

1..818  
 /organism="Homo sapiens"  
 /mol\_type="rRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30409775"  
 /issue\_type="Pooled"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /note="Vector: pDNR-LIB; Site 1: SfiI (ggccattagcc);  
 Site 2: SfiI (ggcgccgcgcgc); Library is oligo-dT primed  
 and directionally cloned. PBMC - Peripheral Blood  
 Mononuclear Cells. RNA was pooled from 3/hour stimulation  
 with PMA adn Ionomycin. 5' and 3' adaptors were used in  
 cloning as follows: 5' adaptor sequence:  
 5'-CAGGCCCATTTATGCC-3' and 3' adaptor sequence:  
 5'-ATTCTAGAGCCGCGGCCGACATG-3' (where B = A,  
 C, or G and N = A, C, G, or T). Average insert size 1.69  
 kb (range 0.70-5.0 kb). 15/15 colonies contained inserts  
 by PCR. This library was enriched for full-length clones  
 and was constructed by Clontech Laboratories (Palo Alto,  
 CA). Note: this is a NIH MGC Library."

## ORIGIN

Query Match 21.3%; Score 446; DB 6; Length 818;

Best Local Similarity 99.6%; Pred. No. 1.7e-233; Indels 0; Gaps 0;

Matches 546; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

425 GGATTGCTGAGCTCTGGCAGCCCTCAGAAAGATTGTTAGTCAAGCAGAGCAAGAGG 484  
 3 GGATTGCTGAGCTCTGGCAGCCCTCAGAAAGATTGTTAGTCAAGCAGAGCAAGAGG 62  
 485 AATATAGTGGCGACCTCCGGTGGCATCCTGATCCTTGAAGAGACGCTTGTAACT 544  
 63 AATATAGTGGCGACCTCCGGTGGCATCCTGATCCTTGAAGAGACGCTTGTAACT 122  
 545 CAGAGCTGCTACATTTGATCTCAAGTGAAGCTTCTGAGATTCAGATTCAGAGAA 604  
 123 CAGAGCTGCTACATTTGATCTCAAGTGAAGCTTCTGAGATTCAGATTCAGAGAA 182  
 605 AGGATTCGAGGTTTGAAGCAAAATGCAAGTGAAGCAACCAATCAATGTTGTAATG 664  
 183 AGGATTCGAGGTTTGAAGCAAAATGCAAGTGAAGCAACCAATCAATGTTGTAATG 242  
 665 AAGACTTGAAGTCTTCACTTACCGTTCGGTACCCCACTCTCAAGAGCTCTTGAA 724  
 243 AAGACTTGAAGTCTTCACTTACCGTTCGGTACCCCACTCTCAAGAGCTCTTGAA 302  
 725 TTCCTGGTTTACCCCGAATATTTACAGTATCTGAGAGAGTCTTGGCAGAGAG 784  
 303 TTCCTGGTTTACCCCGAATATTTACAGTATCTGAGAGAGTCTTGGCAGAGAG 362  
 785 AAGCAAGATCTGATCTGATCTGAGAGATTCAGATTCAGATTCAGATTCAGAGAG 844  
 363 AAGCAAGATCTGATCTGATCTGAGAGATTCAGATTCAGATTCAGATTCAGAGAG 422  
 845 TTCAACTTACAGATCTGATCTGAGAGATTCAGATTCAGATTCAGATTCAGAGAG 904  
 423 TTCAACTTACAGATCTGATCTGAGAGATTCAGATTCAGATTCAGATTCAGAGAG 482  
 905 ATACAGACTTTTCTTACAGCTGAGAGATTCAGATTCAGATTCAGATTCAGAGAG 964  
 483 ATACAGACTTTTCTTACAGCTGAGAGATTCAGATTCAGATTCAGATTCAGAGAG 542  
 965 CTGAGAGTA 972  
 543 CTGAGAGTA 550

## RESULT 15

BI025283

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AUTHORS

REFERENCE

JOURNAL

MEDLINE

PUBMED

COMMENT

TITLE

Shotgun sequencing of the human transcritpome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

10737800

Contact: Simpson A.J.G.

Ludwig Institute for Cancer Research

Rue Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the PABSP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(<http://www.ludwig.org.br/scripts/gethtml2.pl?el=RCS&2=RCS-MT0259-020201-021-G04&3=2001-02-02&4=1>)

Seq primer: puc 18 forward

High quality sequence stop: 78

High quality sequence stop: 590.

Location/Qualifiers

1..591

/organism="Homo sapiens"

/mol\_type="rRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

/clone\_1b="MT0259"

/note="Organ: marrow; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTRS PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

## ORIGIN

Query Match 20.7%; Score 434; DB 4; Length 591;

Best Local Similarity 99.8%; Pred. No. 6.6e-227; Indels 0; Gaps 0;

Matches 484; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

525 GAGGACAGCCTTGAAGTCAAGCTGCTACAGTGAATCTCAAGTGAAGTCTGAG 584  
 102 GAGGACAGCCTTGAAGTCAAGCTGCTACAGTGAATCTCAAGTGAAGTCTGAG 161  
 585 ATTGATGATTCAGAGAAAGATTTCTGAGTTTGAAGCAAAATGCAAGTGAAGCAA 644  
 162 ATTGATGATTCAGAGAAAGATTTCTGAGTTTGAAGCAAAATGCAAGTGAAGCAA 221  
 645 CCAATTCATTTGTTGTAATGAAGACTTGAAGTCTCACTTACCCGTTCCGTTACCCCACT 704  
 222 CCAATTCATTTGTTGTAATGAAGACTTGAAGTCTCACTTACCCGTTCCGTTACCCCACT 281  
 705 CTCACAGGCTCTGGAATATTCCTGTTTACCCCGAATATTTACAGTACATCTGCA 764

Db	282	CTGACAAAGCCTCTCTGAATATTCCTGTTTACCCCGAAATTTACAGGTACATCTGCA	341
Qy	765	GGAATCTCTTGCCAGGAGGAAAGCCAAATATCTGTGACTTTCAGCAGATCCAGTTTTC	824
Db	342	GGAATCTCTTGCCAGGAGGAAAGCCAAATATCTGTGACTTTCAGCAGATCCAGTTTTC	401
Qy	825	AGTCCCAATTTCAAAGGCACTTCACTTACTAGCAATGATGCCATTAACCACTCTGCT	884
Db	402	AGTCCCAATTTCAAAGGCACTTCACTTACTAGCAATGATGCCATTAACCACTCTGCT	461
Qy	885	GATGAATTTGACATTTCAATACAGACTTTTCTATCAGCTGAGATGCTTCAGCCT	944
Db	462	GATGAATTTGACATTTCAATACAGACTTTTCTATCAGCTGAGATGCTTCAGCCT	521
Qy	945	GATCTGCCCTTAACAGTATTTCTGAGTACAAGCCTACTCCAAAGACTGCGCTTGAGA	1004
Db	522	GATCTGCCCTTAACAGTATTTCTGAGTACAAGCCTACTCCAAAGACTGCGCTTGAGA	581
Qy	1005	TAAAA 1009	
Db	582	TAAAA 586	

Search completed: August 27, 2005, 15:58:39  
 Job time : 4539.36 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 19:17:21 ; Search time 5995.75 Seconds  
(without alignments)  
16914.771 Million cell updates/sec

Title: US-09-371-347A-47

Perfect score: 2093  
Sequence: 1 atgaggaggttctgttact.....ttcagatattgtgcataa 2093

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:\*

1: gb\_ba:\*\n2: gb\_hgt:\*\n3: gb\_in:\*\n4: gb\_om:\*\n5: gb\_ov:\*\n6: gb\_pat:\*\n7: gb\_ph:\*\n8: gb\_pl:\*\n9: gb\_pr:\*\n10: gb\_ro:\*\n11: gb\_scs:\*\n12: gb\_sy:\*\n13: gb\_un:\*\n14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1793	85.7	3259	6	AR144976 Sequence
2	1793	85.7	3259	6	AX050463 Sequence
3	1793	85.7	3259	6	AF025794 Homo sapi
4	1793	85.7	3259	6	AF121214 Homo sapi
5	1691	80.8	3241	6	CQ726091 Sequence
6	1640	78.4	3310	9	BC054816 Homo sapi
7	1220	58.3	2933	11	BV177620 sqm95800
8	1220	58.3	2933	11	BV178010 sqm97986
9	386	18.4	330	6	BD077780 5' EST of
10	381	18.2	1353	9	AF121205 Homo sapi
11	330	15.8	109626	9	AC010346 Homo sapi
12	330	15.8	110756	9	AC025174 Homo sapi
13	279	13.3	158199	2	AC022921 Homo sapi
14	279	13.3	167237	2	AC021609 Homo sapi
15	279	13.3	177596	2	AC091945 Homo sapi
16	189	9.0	1156	9	AF121210 Homo sapi
17	188	9.0	1034	9	AF121202S01
18	188	9.0	167237	2	AC021609 Homo sapi
19	188	9.0	177596	2	AC091945 Homo sapi

20	183	8.7	1432	9	FI21202S07	AF121208 Homo sapi
21	161	7.7	158199	2	AC022921	AC022921 Homo sapi
22	158	7.5	1256	9	FI21202S02	AF121203 Homo sapi
23	158	7.5	2475	6	AR454615	AR454615 Sequence
24	158	7.5	2475	6	AX375651	AX375651 Sequence
25	155	7.4	2011	9	FI21202S06	AF121207 Homo sapi
26	146	7.0	2214	9	FI21202S12	AF121213 Homo sapi
27	129	6.2	4506	9	FI21202S01	AF121206 Homo sapi
28	125	6.0	969	9	FI21202S05	AF121206 Homo sapi
29	121	5.8	1119	9	FI21202S10	AF121211 Homo sapi
30	119	5.7	1200	9	FI21202S03	AF121204 Homo sapi
31	78	3.7	243	6	BD04284	BD04284 5' EST and
32	63	3.0	63	6	AX611833	AX611833 Sequence
33	60	2.9	60	6	CO539377	CO539377 Sequence
34	54	2.6	54	6	AX611839	AX611839 Sequence
35	54	2.6	54	6	AX611843	AX611843 Sequence
36	51	2.4	51	6	AX162161	AX162161 Sequence
37	48	2.3	48	6	AX611835	AX611835 Sequence
38	48	2.3	48	6	AX611841	AX611841 Sequence
39	47	2.2	183	6	CO670532	CO670532 Sequence
40	44	2.1	650	9	FI21202S08	AF121209 Homo sapi
41	41	2.0	41	6	AX611845	AX611845 Sequence
42	38	1.8	38	6	AX611837	AX611837 Sequence
43	38	1.8	63	6	AX611834	AX611834 Sequence
44	32	1.5	238720	2	AC095949	AC095949 Ratfus no
45	32	1.5	271339	2	AC131637	AC131637 Ratfus no

## ALIGNMENTS

RESULT 1	AR144976	3259 bp	DNA	linear	PAT 08-AUG-2001
LOCUS	AR144976	Sequence 23 from patent US 6210950.			
DEFINITION	AR144976				
ACCESSION	AR144976.1	GI:15106843			
VERSION					
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 3259)				
AUTHORS	Johnson, W.G. and Stemroos, R.Scott.				
TITLE	Methods for diagnosing, preventing, and treating developmental disorders due to a combination of genetic and environmental factors				
JOURNAL	Patent: US 6210950-A 23 03-Apr-2001;				
FEATURES	Location/Qualifiers				
source	1..3259				
ORIGIN	/organism="unknown"				
	/mol_type="unassigned DNA"				
Query Match	85.7%;	Score 1793;	DB 6;	Length 3259;	
Best Local Similarity	99.8%;	Pred. No. 0;			
Matches 2093;	Conservative	0;	Mismatches	0;	Indels 4; Gaps 1;
QY	1	ATGAGGAGGTTTCTGTACTATATGCTACACGACGAGGCAAGCCATCCGAGAA	60		
DB	80	ATGAGGAGGTTTCTGTACTATATGCTACACGACGAGGCAAGCCATCCGAGAA	139		
QY	61	GAATGTGTAGCAAGCTGTGTATCTTGTGAGATCTTCACTGATTAAGTAA	120		
DB	140	GAATGTGTAGCAAGCTGTGTATCTTGTGAGATCTTCACTGATTAAGTAA	199		
QY	121	TCGATTAAGTATGACCTTAAACCGAAGAGCTCTTGTGTGTTTCTACACG	180		
DB	200	TCGATTAAGTATGACCTTAAACCGAAGAGCTCTTGTGTGTTTCTACACG	259		
QY	181	GGACCGGAGACCCACCCGACACAGCCCGCAAGTTTGAAGAAATACGAACCAACA	240		
DB	260	GGACCGGAGACCCACCCGACACAGCCCGCAAGTTTGAAGAAATACGAACCAACA	319		
QY	241	CTGCCGTTGATTCTTCTCACTGACCTGCGGTATGGGTACTGGGTCTCGTATTCAGAA	300		

Db	1400	CTGCTGCAACATCTTCTCTAACTTCAACCCAGACACATATTCGTGTGAGCAAGCTTCA	1455
Qy	1381	TTTCACCCAGGAAGAGCTTCCATTTTGTCTTCAACATTTGTGAAATTTCTGTCTACTGCCACA	1440
Db	1460	TTTTCACCCAGGAAGAGCTTCCATTTTGTCTTCAACATTTGTGAAATTTCTGTCTACTGCCACA	1519
Qy	1441	ACAGAAGTTCTCGGGAAAGGGAGTATGTACAGGCTGCGCTGGCTTGGTGTGCTTCACTT	1500
Db	1520	ACAGAAGTTCTCGGGAAAGGGAGTATGTACAGGCTGCGCTGGCTTGGTGTGCTTCACTT	1579
Qy	1501	CTTGAGCCAAACATACATATGATGCCATGAGACAGCGGGAAGGCCCTGTGCTTAAGATA	1560
Db	1580	CTTGAGCCAAACATACATATGATGCCATGAGACAGCGGGAAGGCCCTGTGCTTAAGATA	1639
Qy	1561	TTCAATCTTCTCTCGAACAACAATTTCTTTCACCTTACAGATGACCCCTCAATCCCATC	1620
Db	1640	TTCAATCTTCTCTCGAACAACAATTTCTTTCACCTTACAGATGACCCCTCAATCCCATC	1699
Qy	1621	ATTATGTTGGTCTCAGGAACCGGATATGCCCCCTTATTTGGGCTCTACAAACAT---AG	1676
Db	1700	ATTATGTTGGTCTCAGGAACCGGATATGCCCCCTTATTTGGGCTCTACAAACATGAGAG	1755
Qy	1677	AAACTCCAGAACAACAACCAGATGGAATTTTGTGAGCAATGTGGTGTTTTGTGGCTGC	1736
Db	1760	AAACTCCAGAACAACAACCAGATGGAATTTTGTGAGCAATGTGGTGTTTTGTGGCTGC	1819
Qy	1737	AGGCATTAAGATTAAGATTAATCTATTCAGAAAAGAGCTCAGACATTTCTTAAGCATGGG	1786
Db	1820	AGGCATTAAGATTAAGATTAATCTATTCAGAAAAGAGCTCAGACATTTCTTAAGCATGGG	1879
Qy	1797	ATCTTAATCTCATCTTAAAGGTTTCTTCTTCAGAGAATGTCTCTGTTTGGGAGAGAGAACCC	1855
Db	1880	ATCTTAATCTCATCTTAAAGGTTTCTTCTTCAGAGAATGTCTCTGTTTGGGAGAGAGAACCC	1939
Qy	1857	CCAGCAAAAGTATGTACAAACAACATCCAGCTTCATGGCCAGACAGTGGCGAGATCTCTC	1918
Db	1940	CCAGCAAAAGTATGTACAAACAACATCCAGCTTCATGGCCAGACAGTGGCGAGATCTCTC	1999
Qy	1917	CTCCAGAGGAACGGCCATATTTATGTGTGTGAGATGCAAGAAATATGCGCCAAAGATGTA	1976
Db	2000	CTCCAGAGGAACGGCCATATTTATGTGTGTGAGATGCAAGAAATATGCGCCAAAGATGTA	2055
Qy	1977	CATGATGCCCTTGTGCAATATATAGCAAGAAGTTGGAGTTGAAAACTTGAAGCAATG	2033
Db	2060	CATGATGCCCTTGTGCAATATATAGCAAGAAGTTGGAGTTGAAAACTTGAAGCAATG	2119
Qy	2037	AAAACCCCTGGCACTTTTAAAAAGAAAGAAAGCGTACCTTCAGAGTATTTGGCTATTA	2093
Db	2120	AAAACCCCTGGCACTTTTAAAAAGAAAGAAAGCGTACCTTCAGAGTATTTGGCTATTA	2176
RESULT 2			
AX050463			
LOCUS	AX050463	3259 bp	DNA linear PAT 12-JAN-2001
DEFINITION	Sequence 23 from Patent WO0071754.		
ACCESSION	AX050463		
VERSION	AX050463.1	GI:1226668	
KEYWORDS	.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
REFERENCE	1		
AUTHORS	Johnson W.G. and Stenroos, E.S.		
TITLE	Methods for diagnosing, preventing, and treating developmental disorders due to a combination of genetic and environmental factors		
JOURNAL	Patent: WO 0071754-A 23 30-NOV-2000;		
	University of Medicine and Dentistry of New Jersey (US)		
FEATURES	location/Qualifiers		
SOURCE	1..3259		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		



RESULT 3  
AF025794 3259 bp mRNA linear PRI 26-MAR-1998  
LOCUS Homo sapiens methionine synthase reductase (MTRR) mRNA, complete  
DEFINITION  
ACCESSION AF025794  
VERSION AF025794.1 GI:2981302  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 3259)  
Leclerc, D., Wilson, A., Dumas, R., Gafuick, C., Song, D., Watkins, D.,  
Heng, H.H.Q., Rommens, J.M., Scherer, S.W., Rosenblatt, D.S. and  
Gravel, R.A.  
Cloning and mapping of a cDNA for methionine synthase reductase, a  
flavoprotein defective in patients with homocystinuria  
Proc. Natl. Acad. Sci. U.S.A. 95 (6), 3059-3064 (1998)  
MEDLINE  
PUBMED 9501215  
TITLE  
JOURNAL 2 (bases 1 to 3259)  
AUTHORS Leclerc, D.  
REFERENCE  
JOURNAL Direct Submission  
TITLE Submitted (19-SEP-1997) Human Genetics, McGill University -  
AUTHORS Montreal Children's Hospital Research Institute, 4060 Ste-Catherine  
West, Montreal, Que H3Z 2Z3, Canada  
3 (bases 1 to 3259)  
Leclerc, D.  
REFERENCE  
JOURNAL Direct Submission  
TITLE Submitted (12-NOV-1997) Human Genetics, McGill University -  
AUTHORS Montreal Children's Hospital Research Institute, 4060 Ste-Catherine  
West, Montreal, Que H3Z 2Z3, Canada  
REMARK  
FEATURES  
SOURCE  
1..3259  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/map="5p15.2-p15.3"  
1..3259  
/gene="MTRR"  
80..2176  
/gene="MTRR"  
/note="B12 vitamin; cobal complementation group"  
/codon\_start=1  
/product="methionine synthase reductase"  
/protein\_id="AAC39667.1"  
/translation="MRRFLLYVATQGOAKAIAEMCEQAVHGFSAADLHICISBDKY  
DLKETAPLVVVSTGTGDPPTARKFVEIONQTLPVDFPHLRKGLGLDSEYT  
YFNGGKIDIKRLOELGARHFDYDGHADGCVGLIVVEPWIAGLWPALRHPSSNG  
EETSGAPVASPASRTDYKSELHIESVELLRPDSRKSEVILKQNAVNSNG  
VIEDFESSILRSVPLSQASLNTPGIPETYLQVHLQESLQESQSVTSADVQV  
PIKSAVOLTTNDVILKTLVLELIDISNDFSYQPDASVLCPSNDSVOSLRLDLE  
DKREHCYLLIKADTKKKGATLPOHLPAGSLQFIPTWCHEIRAIPKVALRLVOT  
SDSAEKRLLOELCSKOGAAYNRPVACGLDLILAFSPCCPISLIEHLPKLP  
RYSKASSSLPHGKLFHFNIVEPSTATTEVLARKVCTGMAALIVASTLQNTIAS  
HEBSGRKALAKISISPTTSSPHLPDPSPIPIIMVGCTGTAIPGILQREKLQEOH  
PDGNFGAMWLPFGGRHQRDYLFEKELRHLKGLILHLVSSSRDAPVEEERAPAY  
VDNITQHGQVARIILQENGHIVCGDAKNMAKDVALVOIISKEVGEKLEAMKT  
LATLKEKRYLQDIWS"

ORIGIN  
Query Match 85.7%; Score 1793; DB 9; Length 3259;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2093; Conservative 0; Mismatches 0; Indels 4; Gaps 1;  
QY 1 ATGAGAGGTTTCTGTACTATATGCTACACAGAGGAGACAGCAAGGCATCGCGAA 60  
DB 80 ATGAGAGGTTTCTGTACTATATGCTACACAGAGGAGACAGCAAGGCATCGCGAA 139

QY 61 GAAATGTGAGACAGCTGTGTACATGATTTTCTGCAGATCTTCACTGATTAAGTAA 120  
DB 140 GAAATGTGAGACAGCTGTGTACATGATTTTCTGCAGATCTTCACTGATTAAGTAA 199  
QY 121 TCCGATTAAGTATGACCTTAATAACCGGAAGAGCTCCCTGTGTGTGTGTTCTTACACG 180  
DB 200 TCCGATTAAGTATGACCTTAATAACCGGAAGAGCTCCCTGTGTGTGTGTTCTTACACG 259  
QY 181 GGCACCGGAGACCCACCCGACACAGCCCGCAAGTTGTTAAGAAATTAAGAACCAACA 240  
DB 260 GGCACCGGAGACCCACCCGACACAGCCCGCAAGTTGTTAAGAAATTAAGAACCAACA 319  
QY 241 CTGCGGTTGATTTCTTTGCTCACCTGCGGATAGGTTAATGAGTTCTCGGTATTCAGAA 300  
DB 320 CTGCGGTTGATTTCTTTGCTCACCTGCGGATAGGTTAATGAGTTCTCGGTATTCAGAA 379  
QY 301 TACACTACTTTTGGCAATGGGGGGAAGATTAATGATTAACCACTTGAAGCTTGAAGCC 360  
DB 380 TACACTACTTTTGGCAATGGGGGGAAGATTAATGATTAACCACTTGAAGCTTGAAGCC 439  
QY 361 CGGCATTTCTATGACACTGGACATGACAGATGACTGTAGGTTTGAACCTTGTTGAG 420  
DB 440 CGGCATTTCTATGACACTGGACATGACAGATGACTGTAGGTTTGAACCTTGTTGAG 499  
QY 421 CCGTGATTTGCTGACTCTGGCCAGCTCTGAGAACGATTTTNGTCAAGACAGACAA 480  
DB 500 CCGTGATTTGCTGACTCTGGCCAGCTCTGAGAACGATTTTNGTCAAGACAGACAA 559  
QY 481 GAGGATATAGTGGGGGCACTCCCGGNGGATCCCTGACCTTGAAGGACAGCTTGG 540  
DB 560 GAGGATATAGTGGGGGCACTCCCGGNGGATCCCTGACCTTGAAGGACAGCTTGG 619  
QY 541 AAGTCAGAGCTGTACACATTTGATCTCAAGTCGAGCTTCTGAGATTCAGATTCAGGA 600  
DB 620 AAGTCAGAGCTGTACACATTTGATCTCAAGTCGAGCTTCTGAGATTCAGATTCAGGA 679  
QY 601 AGAAGGATTTGAGGTTTGAAGCAAAATGACGTGAACAGCAACCAATCCAAATGTTGTA 660  
DB 680 AGAAGGATTTGAGGTTTGAAGCAAAATGACGTGAACAGCAACCAATCCAAATGTTGTA 729  
QY 661 ATTGAAGCTTTGAGGCTTCACTTACCCGTTGGTACCCCACTTCAAGGCTCTG 720  
DB 740 ATTGAAGCTTTGAGGCTTCACTTACCCGTTGGTACCCCACTTCAAGGCTCTG 799  
QY 721 AATATTCCTGTTTACCCCAAGATATTTTACAGGTACATCTGAGAGAGTCTTGGCCAG 780  
DB 800 AATATTCCTGTTTACCCCAAGATATTTTACAGGTACATCTGAGAGAGTCTTGGCCAG 859  
QY 781 GAGGAAAGCCAGTATCTGTGACTTACAGAGATCCAGTTTTCAGTGCCTTCAAG 840  
DB 860 GAGGAAAGCCAGTATCTGTGACTTACAGAGATCCAGTTTTCAGTGCCTTCAAG 919  
QY 841 GCGTTCACTTACCTTACAGATATGATGCAATTAACCACTGCTGTGATTAATGAGCAT 900  
DB 920 GCGTTCACTTACCTTACAGATATGATGCAATTAACCACTGCTGTGATTAATGAGCAT 979  
QY 901 TCAATATACAGACTTTTCTATCAGCTGTGAGATGCTTCAAGCTGTGATTCGCTTACAGT 960  
DB 980 TCAATATACAGACTTTTCTATCAGCTGTGAGATGCTTCAAGCTGTGATTCGCTTACAGT 1039  
QY 961 GATTGTAGGTACAAAGCTTACCTCAAGATCTGCAAGCTTGAAGATTAAGAGAGCACTGC 1020  
DB 1040 GATTGTAGGTACAAAGCTTACCTCAAGATCTGCAAGCTTGAAGATTAAGAGAGCACTGC 1099  
QY 1021 GTCCCTTTGAAAATAAGGAGACAAAGAAAGAGAGCTTCACTCCCGAGCATATA 1080  
DB 1100 GTCCCTTTGAAAATAAGGAGACAAAGAAAGAGAGCTTCACTCCCGAGCATATA 1159  
QY 1081 CTTGCGGAGATGTTCTCTCAGATTCATTTTATCTGTGTTCTTGAATTCGAGAGAAATTCCT 1140  
DB 1160 CTTGCGGAGATGTTCTCTCAGATTCATTTTATCTGTGTTCTTGAATTCGAGAGAAATTCCT 1219



Query Match 85.7%; Score 1793; DB 9; Length 3291;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2093; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 1 ATGAGAGGTTCTGTACTATATATGCTACAGAGAGGACAGGCAAAAGCCATCGAGAA 60  
DB 112 ATAGAGAGGTTCTGTACTATATGCTACAGAGAGGACAGGCAAAAGCCATCGAGAA 171  
QY 61 GAAATGTGAGCAAGCTGTGTACATGATTTTCTGCAGATCTTCACTGTATTAGTAA 120  
DB 172 GAAATGTGAGCAAGCTGTGTACATGATTTTCTGCAGATCTTCACTGTATTAGTAA 211  
QY 121 TCCGATTAAGTATGACCTTAACAAACCGAAACAGCTCTTGTGTGTGTGTCTACACG 180  
DB 232 TCCGATTAAGTATGACCTTAACAAACCGAAACAGCTCTTGTGTGTGTGTCTACACG 291  
QY 181 GGCACCGAGAACCCAGCACAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACA 240  
DB 292 GGCACCGAGAACCCAGCACAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACA 351  
QY 241 CTGCGGTTGATTTCTTTGCTCACTGCGGTATGGGTTACTGGGCTCGGTGATTAGAA 300  
DB 352 CTGCGGTTGATTTCTTTGCTCACTGCGGTATGGGTTACTGGGCTCGGTGATTAGAA 411  
QY 301 TACACCTACTTTTGCAATGGGGGGAGATTAATTGATTAAGACTTTCAAGCTTTGAGCC 360  
DB 412 TACACCTACTTTTGCAATGGGGGGAGATTAATTGATTAAGACTTTCAAGCTTTGAGCC 471  
QY 361 CGGCATTCTATGACATGACATGAGATGACTGTAGTATTAGAACTTTGTGTGTAG 420  
DB 472 CGGCATTCTATGACATGACATGAGATGACTGTAGTATTAGAACTTTGTGTGTAG 531  
QY 421 CGGTGATTCGTGACCTCTGGCCAGCCCTCAGAAAGATTTAGGTCAAGCAGAGCAA 480  
DB 532 CGGTGATTCGTGACCTCTGGCCAGCCCTCAGAAAGATTTAGGTCAAGCAGAGCAA 591  
QY 481 GAGGAGTAAGTGGCGGACCTCCGGTGGCATCACTGCTATCTTGAGGACAGACCTTGTG 540  
DB 592 GAGGAGTAAGTGGCGGACCTCCGGTGGCATCACTGCTATCTTGAGGAGAGACCTTGTG 651  
QY 541 AAGTCAGAGCTGACATGAAATCTCAAGTGGAGCTTCGAGATTCAGTGTACGAA 600  
DB 652 AAGTCAGAGCTGACATGAAATCTCAAGTGGAGCTTCGAGATTCAGTGTACGAA 711  
QY 601 AGAAAGATTCGTAGGTTTGAAGCAAAATGACGTGAACAGCAACCAATCAAATGTTGTA 660  
DB 712 AGAAAGATTCGTAGGTTTGAAGCAAAATGACGTGAACAGCAACCAATCAAATGTTGTA 771  
QY 661 ATTGAAGACTTTGAGTCTCTCACTTACCCTTCGGTACCCCACTCTCAAGACTCTCTG 720  
DB 772 ATTGAAGACTTTGAGTCTCTCACTTACCCTTCGGTACCCCACTCTCAAGACTCTCTG 831  
QY 721 AATATTCCTGATTAACCCCGAGAAATTTTACAGGTATCTGCAGAGTCTCTTGGCCAG 780  
DB 832 AATATTCCTGATTAACCCCGAGAAATTTTACAGGTATCTGCAGAGTCTCTTGGCCAG 891  
QY 781 GAGGAAAGCCAAAGTATCTGTGACTTCAGCAGATTCAGTCTTTTCAAGTCCAAATTTCAAG 840  
DB 892 GAGGAAAGCCAAAGTATCTGTGACTTCAGCAGATTCAGTCTTTTCAAGTCCAAATTTCAAG 951  
QY 841 GCGATTCACTTACTACGAATGATGCAATAAAACCACTGCTGGTGAATTTGAGACTT 900  
DB 952 GCGATTCACTTACTACGAATGATGCAATAAAACCACTGCTGGTGAATTTGAGACTT 1011  
QY 901 TCAAAATCAGACTTTCTCTATCAGCTGAGATGCTTACGCGTATCTGCGCTTAAAGT 960  
DB 1012 TCAAAATCAGACTTTCTCTATCAGCTGAGATGCTTCTTACGCGTATCTGCGCTTAAAGT 1071  
QY 961 GATTCTGAGGTACAAAGCTTACTCAAAAGCTGCAAGCTTGAAGATTAAGAGAGCACTGC 1020  
DB 1072 GATTCTGAGGTACAAAGCTTACTCAAAAGCTGCAAGCTTGAAGATTAAGAGAGCACTGC 1131  
QY 1021 GTCTTTTGAATAAAGGACAGACAAAGAAAGAGCTTACCTTACCCAGCATATA 1080

DB 1132 GTCTTTTAAATAAAGGACAGACAAAGAAAGAGCTTACCTTACCCAGCATATA 1191  
QY 1081 CCTGCGGAGATGTTCTCTCAGTTCATTTTCTACCTGGTGTCTTGAATCCGAGCAATTCCT 1140  
DB 1192 CCGCGGAGATGTTCTCTCAGTTCATTTTCTACCTGGTGTCTTGAATCCGAGCAATTCCT 1251  
QY 1141 AAAAAGCAATTTTGGAGCCCTTGTGACATTAATACAGTGAACGTCTGAAAAGCCAGG 1200  
DB 1252 AAAAAGCAATTTTGGAGCCCTTGTGACATTAATACAGTGAACGTCTGAAAAGCCAGG 1311  
QY 1201 CTACAGAGCTGTGACATTAACAAAGGGGAGCGATTAATAGCGCTTTGACAGATGCG 1260  
DB 1312 CTACAGAGCTGTGACATTAACAAAGGGGAGCGATTAATAGCGCTTTGACAGATGCG 1371  
QY 1261 TGTGCTGTGTGTGATCTCTCTCGCTTCCCTTCTTCCAGCAGCACACTCAGTCTC 1320  
DB 1372 TGTGCTGTGTGTGATCTCTCTCGCTTCCCTTCTTCCAGCAGCACACTCAGTCTC 1431  
QY 1321 CTGCTGAAATCTTCTTAACTTCAACCCAGACATATTCGTGTGCAAGCTCAAGTTTA 1380  
DB 1432 CTGCTGAAATCTTCTTAACTTCAACCCAGACATATTCGTGTGCAAGCTCAAGTTTA 1491  
QY 1381 TTTCAACCAGAAAGCTCCATTTGTCTCAACATTTGTGGAATTTCTGTACTGTGCA 1440  
DB 1492 TTTCAACCAGAAAGCTCCATTTGTCTCAACATTTGTGGAATTTCTGTACTGTGCA 1551  
QY 1441 ACAGAGTTCTGCGAAGGAGATATGACAGGCTGGCTGCTTGTGTGTGCTTCAAGT 1500  
DB 1552 ACAGAGTTCTGCGAAGGAGATATGACAGGCTGGCTGCTTGTGTGTGCTTCAAGT 1611  
QY 1501 CTTCAGCCAAACATCATGATCCCATGAAGACAGCGGAAAGCCCTGCTCTTAAGATA 1560  
DB 1612 CTTCAGCCAAACATCATGATCCCATGAAGACAGCGGAAAGCCCTGCTCTTAAGATA 1671  
QY 1561 TCCATCTCTCTCGAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
DB 1672 TCCATCTCTCTCGAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1731  
QY 1621 ATATATGAGGATCGAGAAACCGGACATAGCCGCTTATTTGGTTCCTAACAAT---AG 1676  
DB 1732 ATATATGAGGATCGAGAAACCGGACATAGCCGCTTATTTGGTTCCTAACAATGAG 1791  
QY 1677 AAACTCAAGAACACACCCAGATGGAATTTTGAAGCATGTGTGTTTTTGGCTGC 1736  
DB 1792 AAACTCAAGAACACACCCAGATGGAATTTTGAAGCATGTGTGTTTTTGGCTGC 1851  
QY 1737 AGGCAATAGGATAGGATTAATCATTCAGAAAAGACTCAGACATTCCTTAAGCATGG 1796  
DB 1852 AGGCAATAGGATAGGATTAATCATTCAGAAAAGACTCAGACATTCCTTAAGCATGG 1911  
QY 1797 ATCTTAATCATCTAAGGTTTCTTCTCAAGAGATGCTCTGTGGGAGAGAGAACGC 1856  
DB 1912 ATCTTAATCATCTAAGGTTTCTTCTCAAGAGATGCTCTGTGGGAGAGAGAACGC 1971  
QY 1857 CAGCAAAAGTATGACAGACAAACATCAGCTTCATGGCCAGCAGGTGGCAGATCTTC 1916  
DB 1972 CAGCAAAAGTATGACAGACAAACATCAGCTTCATGGCCAGCAGGTGGCAGATCTTC 2031  
QY 1917 CTCCAGAGAAACGCGCATATTTATGTGTGTGAGAGCAAAAGAAATTTGGCCAGAGTGA 1976  
DB 2032 CTCCAGAGAAACGCGCATATTTATGTGTGTGAGAGCAAAAGAAATTTGGCCAGAGTGA 2091  
QY 1977 CATGATGCCCTTGTGCAATTAATTAAGCAAAAGAGTTGAGTTGAABAACTAGAGCATG 2036  
DB 2092 CATGATGCCCTTGTGCAATTAATTAAGCAAAAGAGTTGAGTTGAABAACTAGAGCATG 2151  
QY 2037 AAAACCTGCGCATTTTAAAGAAAGAAACGCTACTCTTCAAGATATTTGTGTCATA 2093  
DB 2152 AAAACCTGCGCATTTTAAAGAAAGAAACGCTACTCTTCAAGATATTTGTGTCATA 2208

RESULT 5



CQ726091 LOCUS CQ726091 3241 bp DNA linear PAT 03-FEB-2004  
DEFINITION Sequence 12025 from Patent WO02068579.  
ACCESSION CQ726091  
VERSION CQ726091.1 GI:42288134  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS 1  
TITLE Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
Kite, such as nucleic acid arrays, comprising a majority of  
humaneons or transcripts, for detecting expression and other uses  
thereof  
Patent: WO 02068579-A 12025 06-SEP-2002;  
PE Corporation (NY) (US)  
FEATURES  
source 1..3241  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 80.8%; Score 1691; DB 6; Length 3241;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 2091; Conservative 0; Mismatches 2; Indels 4; Gaps 1;  
QY 1 ATGAGAGAGTTCTGTACTATATGCTACACAGCAGGACAGGCAAGGCCATCGAGAA 60  
DB 80 ATGAGAGAGTTCTGTACTATATGCTACACAGCAGGACAGGCAAGGCCATCGAGAA 139  
QY 61 GAAATGTGAGCAGAGCTGTGTACATGATTTTCTGCGAGATTTCACTGTATTAGTAA 120  
DB 140 GAAATGTGAGCAGAGCTGTGTACATGATTTTCTGCGAGATTTCACTGTATTAGTAA 199  
QY 121 TCCGATTAATGATGACCTAAACCCGAAACAGCTCTCTGTGTGTGTGTCTACACG 180  
DB 200 TCCGATTAATGATGACCTAAACCCGAAACAGCTCTCTGTGTGTGTGTCTACACG 259  
QY 181 GGCACCGAGAGACCCACCGACACAGCCCGCAAGTTGTTAAGAAATACAGAACAA 240  
DB 260 GGCACCGAGAGACCCACCGACACAGCCCGCAAGTTGTTAAGAAATACAGAACAA 319  
QY 241 CTGCCGCTGATTTCTTTGCTCACTGCGGATGAGTTTCTGCGGCTCGGATTTCA 300  
DB 320 CTGCCGCTGATTTCTTTGCTCACTGCGGATGAGTTTCTGCGGCTCGGATTTCA 379  
QY 301 TACACTACTTTTGGAAATGGGGGGAAGATTAATGATTAACGACTTCAAGAGCTTGA 360  
DB 380 TACACTACTTTTGGAAATGGGGGGAAGATTAATGATTAACGACTTCAAGAGCTTGA 439  
QY 361 CGGCATTTCTATGACACTGACATGACATGACATGACATGACATGACATGACATG 420  
DB 440 CGGCATTTCTATGACACTGACATGACATGACATGACATGACATGACATGACATG 499  
QY 421 CGGTGATTTGCTGAGCTTGTGCGACGCTTCAAGAAAGATTTTATGTTCAAGAGAA 480  
DB 500 CGGTGATTTGCTGAGCTTGTGCGACGCTTCAAGAAAGATTTTATGTTCAAGAGAA 559  
QY 481 GAGGAAATAGTGGCCCACTCCGGGTGGATGACCTGCAATCTTGAAGGACAGACTT 540  
DB 560 GAGGAAATAGTGGCCCACTCCGGGTGGATGACCTGCAATCTTGAAGGACAGACTT 619  
QY 541 AAGTCAGAGCTGTACACATTTGATCTCAAGTCGAGCTTCTGAGATTCGATGATTCAG 600  
DB 620 AAGTCAGAGCTGTACACATTTGATCTCAAGTCGAGCTTCTGAGATTCGATGATTCAG 679  
QY 601 AGAAGAGATTTGAGGTTTGAAGCAAAATGACATGACAGCAACCAATCTCAATGTT 660  
DB 680 AGAAGAGATTTGAGGTTTGAAGCAAAATGACATGACAGCAACCAATCTCAATGTT 739  
QY 661 ATTGAAGACTTTGAGTCTCACTTACCCGTTGGGTACCCCACTCTCAAGAGCTCTG 720

DB 740 ATTGAAGACTTTGAGTCTCACTTACCCGTTGGGTACCCCACTCTCAAGAGCTCTG 799  
QY 721 AATATTTCTGCTTACCCCAAGATATTTTACAGATATCTGACAGAGCTTGTGGCCAG 780  
DB 800 AATATTTCTGCTTACCCCAAGATATTTTACAGATATCTGACAGAGCTTGTGGCCAG 859  
QY 781 GAGAAAGCAGATTTCTGTGACTTACAGATTCAGAGTTTCAAGTGCATTTTCAAG 840  
DB 860 GAGAAAGCAGATTTCTGTGACTTACAGATTCAGAGTTTCAAGTGCATTTTCAAG 919  
QY 841 GAGTTCAACTTACAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 900  
DB 920 GAGTTCAACTTACAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 979  
QY 901 TCAATTCAGACTTTTCTATGAGCTGTGAGATGCTTCAAGGATTCAGGCTTCA 960  
DB 980 TCAATTCAGACTTTTCTATGAGCTGTGAGATGCTTCAAGGATTCAGGCTTCA 1039  
QY 961 GATTCGAGGTACAAAGCTTACTCCAAAGACTGAGCTTGAAGATTAAGAGAGCTGC 1020  
DB 1040 GATTCGAGGTACAAAGCTTACTCCAAAGACTGAGCTTGAAGATTAAGAGAGCTGC 1099  
QY 1021 GTCTTTTGAATAAAGCAGACACAAAGAAAGAGCTTACCTTACCCAGCATATA 1080  
DB 1100 GTCTTTTGAATAAAGCAGACACAAAGAAAGAGCTTACCTTACCCAGCATATA 1159  
QY 1081 CCGCGGAGATTTCTCTCAATTTTCTGCTGCTTCTGCTTCTGCTTCTGCTTCT 1140  
DB 1160 CCGCGGAGATTTCTCTCAATTTTCTGCTGCTTCTGCTTCTGCTTCTGCTTCT 1219  
QY 1141 AAAAGGCAATTTTGGAGCCCTGTGACATTAACAGAGACAGTCTGAAAAGCCAG 1200  
DB 1220 AAAAGGCAATTTTGGAGCCCTGTGACATTAACAGAGACAGTCTGAAAAGCCAG 1279  
QY 1201 CTACAGAGCTGTGAGTAAACAGAGGAGCCGATTAAGCCGCTTTTACAGAGATGC 1260  
DB 1280 CTACAGAGCTGTGAGTAAACAGAGGAGCCGATTAAGCCGCTTTTACAGAGATGC 1339  
QY 1261 TGTGCTGCTTGTGATCTCTCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1320  
DB 1340 TGTGCTGCTTGTGATCTCTCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1399  
QY 1321 CTGCTGGAATCTTCT 1380  
DB 1400 CTGCTGGAATCTTCT 1459  
QY 1381 TTTCAACCCAGAAAGCTCAATTTTGTCTCAATTTGGAATTTCTGTCTACTGAC 1440  
DB 1460 TTTCAACCCAGAAAGCTCAATTTTGTCTCAATTTGGAATTTCTGTCTACTGAC 1519  
QY 1441 ACGAGCTTCTCGGAGGAGATGATGACAGCTGAGCTGCTTGTGCTTCAATTT 1500  
DB 1520 ACGAGCTTCTCGGAGGAGATGATGACAGCTGAGCTGCTTGTGCTTCAATTT 1579  
QY 1501 CTTCAGCCAAATCATCATGATCCCATGGAAGACGAGGGAAGCCCTGCTCTAAGATA 1560  
DB 1580 CTTCAGCCAAATCATCATGATCCCATGGAAGACGAGGGAAGCCCTGCTCTAAGATA 1639  
QY 1561 TCCATCTCTCTGGAACCAATTTTCTCACTTACAGATGACCCCTCAATCCCATC 1620  
DB 1640 TCCATCTCTCTGGAACCAATTTTCTCACTTACAGATGACCCCTCAATCCCATC 1699  
QY 1621 ATTAATGAGGTCTCAAGAAACCGGATAGCCCGGTTATTTGGGTTCTTCAACAT 1676  
DB 1700 ATTAATGAGGTCTCAAGAAACCGGATAGCCCGGTTATTTGGGTTCTTCAACAT 1759  
QY 1677 AAACTCCAAAGCAACACCAAGATGGAATTTTGAAGCAATGAGTGTGTTTGGCTGC 1736  
DB 1760 AAACTCCAAAGCAACACCAAGATGGAATTTTGAAGCAATGAGTGTGTTTGGCTGC 1819  
QY 1737 AGCATTAAGATTAAGGATTAATCTATTCAGAAAAGAGCTCAGCAATTCCTTAAGATG 1796



Db 1820 AGCATATAGAGATAGGATTTATCTATTTCAGAAAAAGCTCAGACATTTCTTAAAGCATGG 1879

Qy 1797 ATCTTAATCATCTAAAGGTTTCTCTCTCAAGAGATCTCTGTGTGGGAGAGAAAGCC 1856

Db 1880 ATCTTAATCATCTAAAGGTTTCTCTCTCAAGAGATCTCTGTGTGGGAGAGAAAGCC 1939

Qy 1857 CCAGCAAGATATGTACAGACACATCTCAGCTTCATGCGCAGAGGTGGCGAATCTTC 1916

Db 1940 CCAGCAAGATATGTACAGACACATCTCAGCTTCATGCGCAGAGGTGGCGAATCTTC 1999

Qy 1917 CTCAGAGAGACGGCCATATTTATGTGTGTGAGATCAAGAAATATGCGCAAGATGTA 1976

Db 2000 CTCAGAGAGACGGCCATATTTATGTGTGTGAGATCAAGAAATATGCGCAAGATGTA 2059

Qy 1977 CATGATGCCCTGTGTGCAATATATAGCAAAAGGTTGAGTGAATAAATAGAACATG 2036

Db 2060 CATGATGCCCTGTGTGCAATATATAGCAAAAGGTTGAGTGAATAAATAGAACATG 2119

Qy 2037 AAAACCTGGCCCTTTAAAGAAAGAAACGCTACTCTTCAAGATATTTGTCATTA 2093

Db 2120 AAAACCTGGCCCTTTAAAGAAAGAAACGCTACTCTTCAAGATATTTGTCATTA 2176

RESULT 6  
BC054816 3310 bp mRNA linear PRI 16-SEP-2003

LOCUS  
DEFINITION  
Homo sapiens 5-methyltetrahydrofolate-homocysteine  
methyltransferase reductase, mRNA (CDNA clone IMAGE:5205285),  
partial cds.

VERSION  
BC054816  
BC054816.1 GI:33392775

KEYWORDS  
Homo sapiens (human)

SOURCE  
Homo sapiens

ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS  
Strasberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D.,  
Altschul, S.F., Zeeberg, B., Bueltow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F.,  
Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
Schnee, T.E., Brownstein, M.J., Urdin, T.B., Tohyuki, S.,  
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,  
McKernan, K.J., Hale, J.A., Gunaratne, P.H., Richards, S.,  
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
Vallalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Fahney, J., Helton, E., Kettaman, M., Madan, A., Rodriguez, S.,  
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
Bretzfeld, Y.S., Krzywicki, M.I., Skalska, U., Smalins, D.B.,  
Schneich, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL  
MEDLINE  
PUBMED  
22388257  
12477932

REFERENCE  
AUTHORS  
Strasberg, R.  
Direct Submission  
Submitted (03-JUL-2003) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,  
USA

REMARK  
COMMENT  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabs-rt@mail.nih.gov](mailto:cgabs-rt@mail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc\\_mgc@hgrl.nih.gov](mailto:nisc_mgc@hgrl.nih.gov)

Akhter, N., Ayala, K., Beckstrom-Stenberg, S.M., Benjamin, B.,  
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,  
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,  
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,  
Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCluskey, J.C.,  
McDowell, J., Pearson, R., Stantrop, S., Thomas, P.J., Touchman, J.W.,  
Toung, A., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LNL at: <http://lml.gov>  
Series: IRAK Plate: 115 Row: d Column: 11  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 4505278.

FEATURES  
source  
1..3310  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5205285"  
/tissue\_type="Lung, spleen, fetal, pooled"  
/clone\_id="N1H\_MGC\_122"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
<1..3310  
/gene="MTRR"  
/note="synonym: MSR"  
/db\_xref="LocusID:4552"  
/db\_xref="MIM:602568"  
<1..2149  
/gene="MTRR"  
/codon\_start=2  
/product="MTRR protein"  
/protein\_id="AAH54816.1"  
/db\_xref="GI:33392776"  
/db\_xref="LocusID:4552"  
/db\_xref="MIM:602568"  
/translation="PVARLRAGFTYTCLEWRRLILATQGGAKATAEICEQAV  
VGFPSADLHCIESSDRTDLKTEAPLVAVSTGTGDPDTAKFVKEIONQTLVDF  
FAGLRVGLGLDSEYTFPCNGCKIDKLDELGAHFYDTGHADCGLELVDEWI  
GLWPLRGLHGFSSRQGEISGALPVASPASSRTDLYKSELHIESQVLELRDDEGR  
KQSEVTKQAVNSNGSNVVIDPESLTSVSPPLSASINIKLPEYLOVHQLQESIG  
QRESOVSTASADPVFPVPISSKAVOLTNDIAIKTTLVLELDISNTPSYOGDAFSVIC  
PNSDSVQGLDRLQEDREHRCVLIKTRDKKATTPQHTFAGCSQPTFTWGLE  
ITAIIPKATLRALVDYTSDAERRLQELCSKQADYRFVADACLDLILAFPS  
CQPPISLLEHLPKQPRPYSCASSLFFPGKLFVFNVEFLSTATTVELLRGCTG  
WALVAVSLTEQNIHSHEDSGKALAPKISIRPTNSPHLDDEPIPIVMGPGTGI  
APFGLGRLKQLEQHPDNGFGKMLFPGCHRKDDYFRLRLHFLKGLITHLKV  
SFSRDAVPEEERAPKAYVDNIQIHQGVARILQENGIHYCGDAKNAKQVHDALV  
QISXSEVYBKLEAKMTTLATLKEBRYLDINS"  
59..493  
/gene="MTRR"  
/note="flavodoxin; Region: Flavodoxin"  
/db\_xref="CDD:pfam0258"  
848..2143  
/gene="MTRR"  
/note="CysU; Region: Sulfite reductase, alpha subunit  
(flavo-protein) (Inorganic ion transport and metabolism)"  
/db\_xref="CDD:COG30369"

ORIGIN  
Query Match 78.4%; Score 1640; DB 9; Length 3310;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 2090; Conservative 0; Mismatches 3; Indels 4; Gaps 1;

Qy 1 ATGAGAGGTTTCTGTATCATATGCTACACAGCAGGACAGGCAAGGCCATCGCAGAA 60  
Db 53 ATGAGAGGTTTCTGTATCATATGCTACACAGCAGGACAGGCAAGGCCATCGCAGAA 112

61 GAAATGTGAGCAAGCTGTGTGACATGATTTTCTGAGATCTTCACTGTATTAAGTGA 120  
113 GAAATGTGAGCAAGCTGTGTGACATGATTTTCTGAGATCTTCACTGTATTAAGTGA 172  
121 TCCGATTAAGTATGACCTTAATAAACCGAAGAGCTCTTGTGTGTGTGTCTACACG 180  
173 TCCGATTAAGTATGACCTTAATAAACCGAAGAGCTCTTGTGTGTGTGTCTACACG 232  
181 GGCACCGGAGACCCGACACAGCCCGCAAGTTGTTAAGAAATACGAACCAACA 240  
233 GGCACCGGAGACCCGACACAGCCCGCAAGTTGTTAAGAAATACGAACCAACA 292  
241 CTGCGGATGATTTCTTGTCTCACTGCGGTATGAGTTACTGGGTCTGCGGTATTCAGA 300  
293 CTGCGGATGATTTCTTGTCTCACTGCGGTATGAGTTACTGGGTCTGCGGTATTCAGA 352  
301 TACACCTACTTTTGCAGTGGGGGGAAGATTAATGATTAACGACTTCAAGAGCTTGAGCC 360  
353 TACACCTACTTTTGCAGTGGGGGGAAGATTAATGATTAACGACTTCAAGAGCTTGAGCC 412  
361 CCGCATTTCTATGACATGACATGACATGACATGACATGACATGACATGACATGACAT 420  
413 CCGCATTTCTATGACATGACATGACATGACATGACATGACATGACATGACATGACAT 472  
421 CCGTGAATGCTGGAATCTGCGCAGCCCTCAGAAAGCATTTTAAAGTCAAGCAGAGCA 480  
473 CCGTGAATGCTGGAATCTGCGCAGCCCTCAGAAAGCATTTTAAAGTCAAGCAGAGCA 532  
481 GAGAGATTAAGTGCAGCATCTCCGATGAGATCACTGCATCTTGAAGCAGACCTTGAG 540  
533 GAGAGATTAAGTGCAGCATCTCCGATGAGATCACTGCATCTTGAAGCAGACCTTGAG 592  
541 AAGTCAGAGCTGTACACATTTGAATCTCAAGTCAGACTTCTGAGATTCATGATTCAGA 600  
593 AAGTCAGAGCTGTACACATTTGAATCTCAAGTCAGACTTCTGAGATTCATGATTCAGA 652  
601 AAAAAAGATTTCTGAGGTTTGAAGCAAAATGACAGTAACAGCAACCAATCCAAATGTTGA 660  
653 AAAAAAGATTTCTGAGGTTTGAAGCAAAATGACAGTAACAGCAACCAATCCAAATGTTGA 712  
661 ATTGAAGACTTTGAGTCTCACTTACCCGTCGGTACCCCACTCTCAAGAGCTCTGAG 720  
713 ATTGAAGACTTTGAGTCTCACTTACCCGTCGGTACCCCACTCTCAAGAGCTCTGAG 772  
721 AATATTCCTGTTTACCCCAAGAAATTTACAGATCACTGAGAGAGTCTCTGAGCCAG 780  
773 AATATTCCTGTTTACCCCAAGAAATTTACAGATCACTGAGAGAGTCTCTGAGCCAG 832  
781 GAGGAAAGCCAGATCTGTGACTTTCAGAGATCCAGTTTTCAGATTCAGATTCAGAA 840  
833 GAGGAAAGCCAGATCTGTGACTTTCAGAGATCCAGTTTTCAGATTCAGATTCAGAA 892  
841 GAGGATCACTTACTAGATGATGACATTAATAACCACTCTGCTGTAGATTCAGATTC 900  
893 GAGGATCACTTACTAGATGATGACATTAATAACCACTCTGCTGTAGATTCAGATTC 952  
901 TCAATATACAGACTTTCTATACAGCTGTGAGATGCTTCAAGCTGTGAGATTCAGAT 960  
953 TCAATATACAGACTTTCTATACAGCTGTGAGATGCTTCAAGCTGTGAGATTCAGAT 1012  
961 GATTCGAGATCAAGAGCTTCTCAAGATCTGAGATTCAGATTCAGATTCAGATTCAG 1020  
1013 GATTCGAGATCAAGAGCTTCTCAAGATCTGAGATTCAGATTCAGATTCAGATTCAG 1072  
1021 GTCCTTTTGAATAAAGGACACAAAGAAAGAAAGAGCTTACCCGACATATTA 1080  
1073 GTCCTTTTGAATAAAGGACACAAAGAAAGAAAGAGCTTACCCGACATATTA 1132  
1081 CTTGCGGAGATGTTCTCTCAGATTCATTTTACCTGAGTCTTGAATTCGAGCAATTCCT 1140  
1133 CTTGCGGAGATGTTCTCTCAGATTCATTTTACCTGAGTCTTGAATTCGAGCAATTCCT 1192

1141 AAAAAAGCATTTTTCGAGCCCTTGTGACATTAACAGTGAAGTGTGAAAAAGCCAGG 1200  
1193 AAAAAAGCATTTTTCGAGCCCTTGTGACATTAACAGTGAAGTGTGAAAAAGCCAGG 1252  
1201 CTACAGAGCTGTGACATTAACAGAGGCGAGCCGATTAATAGCCGCTTGTAGAGATGCG 1260  
1253 CTACAGAGCTGTGACATTAACAGAGGCGAGCCGATTAATAGCCGCTTGTAGAGATGCG 1312  
1261 TGTGCTGCTTGTGATCTCTCTGAGCTTCCCTTCTGAGCAGCAGCATCAGTCTC 1320  
1313 TGTGCTGCTTGTGATCTCTCTGAGCTTCCCTTCTGAGCAGCAGCATCAGTCTC 1372  
1321 CTGCTGAAACATCTTCTTAACCTTCAACCCAGACCATTAATGCTGACAGCTCAAGTTTA 1380  
1373 CTGCTGAAACATCTTCTTAACCTTCAACCCAGACCATTAATGCTGACAGCTCAAGTTTA 1432  
1381 TTTCAACCCAGAAAGCTCATTTTGTCTTCAACATTTGAGAAATTTCTGTCTACGCCACA 1440  
1433 TTTCAACCCAGAAAGCTCATTTTGTCTTCAACATTTGAGAAATTTCTGTCTACGCCACA 1492  
1441 AAGAGGTTCTGCGGAAAGGAGATGTAAGGCTGAGCTGAGCTTGTGAGTTCAGTT 1500  
1493 AAGAGGTTCTGCGGAAAGGAGATGTAAGGCTGAGCTGAGCTTGTGAGTTCAGTT 1552  
1501 CTTCAGCCAAACATCATGATCCCATGAGACACAGCGGAAAGCCCTGCTCTTAAGATA 1560  
1553 CTTCAGCCAAACATCATGATCCCATGAGACACAGCGGAAAGCCCTGCTCTTAAGATA 1612  
1561 TCCATCTCTCTGAAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
1613 TCCATCTCTCTGAAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1672  
1621 AATATGTGAGTTCAGAAACCGGCAATGCCCCGTTTATTTGGTCTCAACAAT----AG 1676  
1673 AATATGTGAGTTCAGAAACCGGCAATGCCCCGTTTATTTGGTCTCAACAATGAGAG 1732  
1677 AATCTCAAGAACCAACACCCAGATGAAATTTTGAAGCATGTGTTGTTTTTGGCTGC 1736  
1733 AATCTCAAGAACCAACACCCAGATGAAATTTTGAAGCATGTGTTGTTTTTGGCTGC 1792  
1737 AAGCATTAAGATAGGATTAATCTATTCAGAAAGAGCTCAGACATTTCTTAAGCATGG 1796  
1793 AAGCATTAAGATAGGATTAATCTATTCAGAAAGAGCTCAGACATTTCTTAAGCATGG 1852  
1797 ATCTTAACATCTTAAGGTTTCTTCTCAAGATGCTCTCTGTTGGGAGAGAGAGCC 1856  
1853 ATCTTAACATCTTAAGGTTTCTTCTCAAGATGCTCTCTGTTGGGAGAGAGAGCC 1912  
1857 CCAGCAAGATGATTAAGAACCAATCCAGCTTCAATGACCAGAGTGGCAGAAATCTCTC 1916  
1913 CCAGCAAGATGATTAAGAACCAATCCAGCTTCAATGACCAGAGTGGCAGAAATCTCTC 1972  
1917 CTTCAGAGAACGAGCATTTTATGTGTGTGAGATGCAAGAAATATGSCAAGATGTA 1976  
1973 CTTCAGAGAACGAGCATTTTATGTGTGTGAGATGCAAGAAATATGSCAAGATGTA 2032  
1977 CATGATCCCTTGTGCAATTAATTAAGCAAGAGTGTGAGATGCAAGAAATTAAGCAAG 2036  
2033 CATGATCCCTTGTGCAATTAATTAAGCAAGAGTGTGAGATGCAAGAAATTAAGCAAG 2092  
2037 AAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2093  
2093 AAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2149

RESULT 7  
BV177620/c 2933 bp DNA linear STS 10-JUN-2004  
LOCUS BV177620  
DEFINITION sqm95800 Human DNA (Sequenc) Homo sapiens STS genomic, sequence  
ACCESSION BV177620  
VERSION BV177620.1 GI:48013757  
KEYWORDS STS.

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 2933)  
 AUTHORS Nelson,R.M., Marnell, G., Kammerer, S., Hoyal, C.R., Shi, M.M.,  
 Cantor, C.R. and Braun, A.  
 TITLE Large-Scale Validation of Single Nucleotide Polymorphisms in Gene  
 Regions  
 JOURNAL Genome Res. (2004) In press  
 COMMENT  
 Contact: Andreas Braun  
 Pharmaceuticals division  
 Sequenom, Inc.  
 3595 John Hopkins Court, San Diego, CA 92121, USA  
 Tel: 18582029018  
 Fax: 18582029020  
 Email: abraun@sequenom.com  
 Primer A: No primer sequence submitted  
 Primer B: No primer sequence submitted  
 STS size: 2933.  
 FEATURES  
 source  
 Location/Qualifiers  
 1. 2933  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone\_id="Human DNA (Sequenom)"  
 <1..2933  
 ORIGIN  
 STS  
 Query Match 58.3%; Score 1220; DB 11; Length 2933;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 315 CAATGGGGGAGATTTGATTAAGACATTCAGAGCTTGGAGCCCGGATTTCTATGA 374  
 Db 2846 CAATGGGGGAGATTTGATTAAGACATTCAGAGCTTGGAGCCCGGATTTCTATGA 2787  
 QY 375 CACTGACATGACAGATGCTGTAGGTTTGAACCTTGAGTGGCCGTGATGCTGG 434  
 Db 2786 CACTGACATGACAGATGCTGTAGGTTTGAACCTTGAGTGGCCGTGATGCTGG 2727  
 QY 435 ACTTGGCCAGCCTTCAGAAAGCATTTTGGTCAAGACAGACAAAGAGATTAAGTGG 494  
 Db 2726 ACTTGGCCAGCCTTCAGAAAGCATTTTGGTCAAGACAGACAAAGAGATTAAGTGG 2667  
 QY 495 CGCATCCCGGTGGCATCACTGCATCCTTGAGAGACAGACCTTGTGAAGTCAGAGCTGCT 554  
 Db 2666 CGCATCCCGGTGGCATCACTGCATCCTTGAGAGACAGACCTTGTGAAGTCAGAGCTGCT 2607  
 QY 555 ACACATTTGAATCTCAAGTCAGAGCTTCTGAGATTCGATTCAGAGAAAGAGATTCCTGA 614  
 Db 2606 ACACATTTGAATCTCAAGTCAGAGCTTCTGAGATTCGATTCAGAGAAAGAGATTCCTGA 2547  
 QY 615 GGTTTTGAAGAAATGATGAAACAGAACCAATCCATGTTTGAATTGAAGCTTTGA 674  
 Db 2546 GGTTTTGAAGAAATGATGAAACAGAACCAATCCATGTTTGAATTGAAGCTTTGA 2487  
 QY 675 GTCTCACTTACCGGTCGGTACCCCACTCTCAAGAGCTCTGTAATTTCTGGTTT 734  
 Db 2486 GTCTCACTTACCGGTCGGTACCCCACTCTCAAGAGCTCTGTAATTTCTGGTTT 2427  
 QY 735 ACCCCAGAAATATTACAGGTACATCTGACAGAGTCTCTTGCCAGAGAGAAAGCCAGT 794  
 Db 2426 ACCCCAGAAATATTACAGGTACATCTGACAGAGTCTCTTGCCAGAGAGAAAGCCAGT 2367  
 QY 795 ATCTGTGATCTTACAGAGATTCAGTTTTCAGATGCAATTTCAAGGAGTTCACTTAC 854  
 Db 2366 ATCTGTGATCTTACAGAGATTCAGTTTTCAGATGCAATTTCAAGGAGTTCACTTAC 2307  
 QY 855 TAGGATGATGCGCATTAAGAAACCACTGCGGTGATGAATTTGACATTTCAATTAAGCTT 914  
 Db 2306 TAGGATGATGCGCATTAAGAAACCACTGCGGTGATGAATTTGACATTTCAATTAAGCTT 2247

QY 915 TTCTTATCAGCCTGAGATGCTTTCAGCGTATCTGCTCTTAAAGTATTTGAGGTACA 974  
 Db 2246 TTCTTATCAGCCTGAGATGCTTTCAGCGTATCTGCTCTTAAAGTATTTGAGGTACA 2187  
 QY 975 AAGCCATCTCCAAAGCTGACCTTGAAGATTAAGAGAGCACTGGTCTTTTGAAGAT 1034  
 Db 2186 AAGCCATCTCCAAAGCTGACCTTGAAGATTAAGAGAGCACTGGTCTTTTGAAGAT 2127  
 QY 1035 AAAGGACAGACAAAGAAAGAAAGAGAGCTTACCTCCAGCATATACCTGCGGATGTTT 1094  
 Db 2126 AAAGGACAGACAAAGAAAGAAAGAGAGCTTACCTCCAGCATATACCTGCGGATGTTT 2067  
 QY 1095 TTCTCAATCTTTTACCTGCTGTTTGAATTCGAGCAATTTCTTAAAGGCAATTTT 1154  
 Db 2066 TTCTCAATCTTTTACCTGCTGTTTGAATTCGAGCAATTTCTTAAAGGCAATTTT 2007  
 QY 1155 GCGAGCCCTTGGGACTATACAGTACAGTGTGTAAGGCGAGGCTACAGAGCTGGT 1214  
 Db 2006 GCGAGCCCTTGGGACTATACAGTACAGTGTGTAAGGCGAGGCTACAGAGCTGGT 1947  
 QY 1215 CAGTAAACAAGGGGACGCGATTTATAGCCGCTTGTGACAGATGCTGTGCTGTGTT 1274  
 Db 1946 CAGTAAACAAGGGGACGCGATTTATAGCCGCTTGTGACAGATGCTGTGCTGTGTT 1887  
 QY 1275 GGAATCTCTCTGCTGCTTTCCTTCTTTCGACGCCACCACTGACTCTGCTGGAACATCT 1334  
 Db 1886 GGAATCTCTCTGCTGCTTTCCTTCTTTCGACGCCACCACTGACTCTGCTGGAACATCT 1827  
 QY 1335 TCTTAAACCTCAACCCAGACCAATTCGTTGTCAGAGCTCAAGTTTATTCACCCAGAGAA 1394  
 Db 1826 TCTTAAACCTCAACCCAGACCAATTCGTTGTCAGAGCTCAAGTTTATTCACCCAGAGAA 1767  
 QY 1395 GCTCAATTTTGTCTTCAACATTTGTGAAATTTCTGTACTGCCCCAAGAGGTTCTGG 1454  
 Db 1766 GCTCAATTTTGTCTTCAACATTTGTGAAATTTCTGTACTGCCCCAAGAGGTTCTGG 1707  
 QY 1455 GAAAGGATGATGTACAGGCTGCTGCTGCTTGTGTTGCTTCAAGTTCTTCAAGCAAACT 1514  
 Db 1706 GAAAGGATGATGTACAGGCTGCTGCTGCTTGTGTTGCTTCAAGTTCTTCAAGCAAACT 1647  
 QY 1515 ACATGATCCCATGAAGACA 1534  
 Db 1646 ACATGATCCCATGAAGACA 1627  
 RESULT 8  
 LOCUS BV178010/c 2933 bp DNA linear STS 10-JUN-2004  
 DEFINITION sqm97986 Human DNA (Sequenom) Homo sapiens STS genomic, sequence  
 tagged site.  
 ACCESSION BV178010  
 VERSION BV178010.1 GI:48014252  
 KEYWORDS STS.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 2933)  
 AUTHORS Nelson,R.M., Marnell, G., Kammerer, S., Hoyal, C.R., Shi, M.M.,  
 Cantor, C.R. and Braun, A.  
 TITLE Large-Scale Validation of Single Nucleotide Polymorphisms in Gene  
 Regions  
 JOURNAL Genome Res. (2004) In press  
 COMMENT  
 Contact: Andreas Braun  
 Pharmaceuticals division  
 Sequenom, Inc.  
 3595 John Hopkins Court, San Diego, CA 92121, USA  
 Tel: 18582029018  
 Fax: 18582029020  
 Email: abraun@sequenom.com  
 Primer A: No primer sequence submitted

Primer B: No primer sequence submitted  
STS size: 2933.  
Location/Qualifiers  
1..2933  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone\_1ib="Human DNA (Sequencem)"  
<1..>2933

STS  
ORIGIN

Query Match 58.3%; Score 1220; DB 11; Length 2933;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 315 CAATGGGGGAGATTAATGATTAAGCACTTCAAGGCTTGAAGCCCGCATTTCTATGA 374
DB 2846 CAATGGGGGAGATTAATGATTAAGCACTTCAAGGCTTGAAGCCCGCATTTCTATGA 2787
QY 375 CACTGGACATGAGATGACCTGTGTAGTTTGAACCTTGTGTGAGCCGTGATGCTGG 434
DB 2786 CACTGGACATGAGATGACCTGTGTAGTTTGAACCTTGTGTGAGCCGTGATGCTGG 2727
QY 435 ACTCTGGCCAGCCCTCAGAAAGCAATTTAGTCAAGCAGAGACAAGAGATTAAGTGG 494
DB 2726 ACTCTGGCCAGCCCTCAGAAAGCAATTTAGTCAAGCAGAGACAAGAGATTAAGTGG 2667
QY 495 CGCATCCCGGTGGCATCACTGCACTTGAAGACAAGACCTTGTGAAGTCAAGCTGCT 554
DB 2666 CGCATCCCGGTGGCATCACTGCACTTGAAGACAAGACCTTGTGAAGTCAAGCTGCT 2607
QY 555 ACACATTAATCAATCAAGTGAAGCTTGAAGTTCAGATTCAGAGAAAGGATTTCTGA 614
DB 2606 ACACATTAATCAATCAAGTGAAGCTTGAAGTTCAGATTCAGAGAAAGGATTTCTGA 2547
QY 615 GGTTTTGAAGCAAAATGACAGTGAACCAACCAATGTTGTAATTGAAGACTTTGA 674
DB 2546 GGTTTTGAAGCAAAATGACAGTGAACCAACCAATGTTGTAATTGAAGACTTTGA 2487
QY 675 GTCTCAGCTTACCCGTTGGGTACCCCACTCTCAAGCCCTCTGTAATATTCCTGTTT 734
DB 2486 GTCTCAGCTTACCCGTTGGGTACCCCACTCTCAAGCCCTCTGTAATATTCCTGTTT 2427
QY 735 ACCCCCAAAATATTTTCAAGTATATCTGACAGAGTCTTTGGCCAGAGAAAGCCAACT 794
DB 2426 ACCCCCAAAATATTTTCAAGTATATCTGACAGAGTCTTTGGCCAGAGAAAGCCAACT 2367
QY 795 ATCTGACCTTCAAGAGATCAAGTTTTCAGAGCCAAATTTCAAGGCAAGTTCAACTTAC 854
DB 2366 ATCTGACCTTCAAGAGATCAAGTTTTCAGAGCCAAATTTCAAGGCAAGTTCAACTTAC 2307
QY 855 TACGAATGATGCCATTAATAACACTGTGCTGTGATGAATTTGACATTTCAATTAACAGACTT 914
DB 2306 TACGAATGATGCCATTAATAACACTGTGCTGTGATGAATTTGACATTTCAATTAACAGACTT 2247
QY 915 TTCTTATCAAGCTTCAAGATTCCTTCAAGCTGATCTGCTTAACAGATTTCTGAGGTACA 974
DB 2246 TTCTTATCAAGCTTCAAGATTCCTTCAAGCTGATCTGCTTAACAGATTTCTGAGGTACA 2187
QY 975 AAGCCTACTCTCAAAAGCTGACGCTTGAATAAAGAGAGACCTGCGTCCCTTTTGAATAAT 1034
DB 2186 AAGCCTACTCTCAAAAGCTGACGCTTGAATAAAGAGAGACCTGCGTCCCTTTTGAATAAT 2127
QY 1035 AAAGGACACACAAAGAAAGAAAGAGCTACCTTACCAGACATATACTGCGGAGATGTC 1094
DB 2126 AAAGGACACACAAAGAAAGAAAGAGCTACCTTACCAGACATATACTGCGGAGATGTC 2067
QY 1095 TCTCCAGTTCAATTTTAACTGTGTCTTGAATAACAGCAATTCCTTAAAGGCAATTTT 1154
DB 2066 TCTCCAGTTCAATTTTAACTGTGTCTTGAATAACAGCAATTCCTTAAAGGCAATTTT 2007
QY 1155 GCGAGCCCTTGTGACATATCAAGTGAAGTGTGAAAAGCGACGCTACAGAGCTGTG 1214
```

```
DB 2006 GCGAGCCCTTGTGACATATCAAGTGAAGTGTGAAAAGCGACGCTACAGAGCTGTG 1947
QY 1215 CAGTAAACAAGGGGAGGAGCGGATTAATAGCGGCTTGTGAAGATGCTGTGCTGTT 1274
DB 1946 CAGTAAACAAGGGGAGGAGCGGATTAATAGCGGCTTGTGAAGATGCTGTGCTGTT 1887
QY 1275 GGATCTCTCTCTGCTGCTTCCCTTCTTCCAGAGCAGCAGCTCACTCTCTGCTGAACATCT 1334
DB 1886 GGATCTCTCTCTGCTGCTTCCCTTCTTCCAGAGCAGCAGCTCACTCTCTGCTGAACATCT 1827
QY 1335 TCTTAACTTCAACCCAGACCATATTCGTGTGCAAGCTCAAGTTATTTACCCAGGAAA 1394
DB 1826 TCTTAACTTCAACCCAGACCATATTCGTGTGCAAGCTCAAGTTATTTACCCAGGAAA 1767
QY 1395 GCTCAATTTTGTCTCAACATTTGTGAAATTTGTGCTACTGTCACAAACAGAGTTCTGG 1454
DB 1766 GCTCAATTTTGTCTCAACATTTGTGAAATTTGTGCTACTGTCACAAACAGAGTTCTGG 1707
QY 1455 GAAGGAGATATGATACAGGCTGAGCTGCTGCTTGTGTGCTTCAAGTTCTTCAACCAACAT 1514
DB 1706 GAAGGAGATATGATACAGGCTGAGCTGCTGCTTGTGTGCTTCAAGTTCTTCAACCAACAT 1647
QY 1515 ACATGCATCCCATGAAGACA 1534
DB 1646 ACATGCATCCCATGAAGACA 1627
```

RESULT 9  
BD077780 390 bp DNA linear PAT 27-AUG-2002  
LOCUS  
DEFINITION 5'EST of secretory protein in brain.  
ACCESSION BD077780  
VERSION BD077780.1 GI:22623383  
KEYWORDS JP 2001512015-A/65.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 390)  
AUTHORS Edwards,J.B.D.M., Duclert,A. and Lacroix,B.  
TITLE 5'EST of secretory protein in brain  
JOURNAL Patent: JP 2001512015-A 65 21-AUG-2001;  
GENSET

COMMENT OS Homo sapiens (human)  
PN JP 2001512015-A/65  
PD 21-AUG-2001  
PF 31-JUL-1998 JP 2000505293  
PR 01-AUG-1997 US 08/905223  
PI JEAN BAPTISTE DUMAS MILNE EDWARDS,AYMERIC DUCLEERT,BRUNO PI  
LACROIX  
PC C12N15/09,C07K14/47,C12M1/00,C12P21/02,C12O1/68,C12N15/00 CC  
Von Heijne matrix  
CC score 6.9  
CC seq SLSLASHSHSVSC/SN  
FH Key Location/Qualifiers  
FT sig\_peptide 289..357.  
Location/Qualifiers  
1..390  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

## ORIGIN

Query Match 18.4%; Score 386; DB 6; Length 390;  
Best Local Similarity 100.0%; Pred. No. 8.7e-199;  
Matches 386; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 970 GTACAAAGCTTCTCAAGATGACGCTTGAAGATTAAGAGAGACGCGCTCTTTG 1029
DB 3 GTACAAAGCTTCTCAAGATGACGCTTGAAGATTAAGAGAGACGCGCTCTTTG 62
QY 1030 AAAATTAAGGACACAAAGAAAGAGAGCTACCTTACCAGATTAATCTGCGGGA 1089
```

Db 63 AAAATAAGGACAGACAAAGAAAGAGCTACCTTACCACGATATACCTGCGGGA 122  
QY 1090 TGTCTCTCAGTTCATTTTACCTGTGTCTTGAATCCGAGCAATTCCTAAAGGCA 1149  
Db 123 TGTCTCTCAGTTCATTTTACCTGTGTCTTGAATCCGAGCAATTCCTAAAGGCA 182  
QY 1150 TTTTGGAGACCTTGTGAGCTATACGAGACAGTGTCTGAAAAAGCGCAGCTACAGAG 1209  
Db 183 TTTTGGAGACCTTGTGAGCTATACGAGACAGTGTCTGAAAAAGCGCAGCTACAGAG 242  
QY 1210 CTGTGACGTAAACAAAGGGGAGCGAGTATATAGCCGCTTTGTACAGATGCTGTGCTGC 1269  
Db 243 CTGTGACGTAAACAAAGGGGAGCGAGTATATAGCCGCTTTGTACAGATGCTGTGCTGC 302  
QY 1270 TTGTGGATCTCTCTCTGCTGCTTTCCTTCTTCCAGCCAGCACTCACTCTCTGCTGAA 1329  
Db 303 TTGTGGATCTCTCTCTGCTGCTTTCCTTCTTCCAGCCAGCACTCACTCTCTGCTGAA 362  
QY 1330 CATCTTCTTAACTTCAACCCAGACC 1355  
Db 363 CATCTTCTTAACTTCAACCCAGACC 388

RESULT 10  
F121202804 1353 bp DNA linear PRI 14-DEC-1999  
LOCUS F121202804 Homo sapiens methionine synthase reductase (MTRR) gene, exon 5.  
ACCESSION AF121205  
VERSION AF121205.1 GI:6572530  
KEYWORDS 4 of 12  
SEGMENT Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1353)  
AUTHORS Leclerc,D., Odievre,M., Wu,Q., Wilson,A., Huizenga,J.J., Rozen,R.,  
Scheyer,S.W. and Gravel,R.A.  
TITLE Molecular cloning, expression and physical mapping of the human  
methionine synthase reductase gene  
JOURNAL Gene 240 (1), 75-88 (1999)  
MEDLINE 20033550  
PUBMED 10564814  
REFERENCE 2 (bases 1 to 1353)  
AUTHORS Leclerc,D.  
TITLE Direct Submission  
JOURNAL Submitted (20-JAN-1999) Human Genetics, Montreal Children's  
Hospital, 4060 Ste-Catherine West, Montreal, Quebec H3Z 2Z3, Canada

FEATURES  
source  
1. .1353  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/map="5p15.2-p15.3; between W11755 and D5S1957"  
358. .736  
/gene="MTRR"  
/number=5

## ORIGIN

Query Match 18.2%; Score 381; DB 9; Length 1353;  
Best Local Similarity 100.0%; Pred. No. 4.2e-196;  
Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 401 GTTTAGAACTTGTGTGAGCCGTGATTTGCTGAGCTCTGGCCAGCCCTCAGAAAGCATT 460  
Db 357 GTTTAGAACTTGTGTGAGCCGTGATTTGCTGAGCTCTGGCCAGCCCTCAGAAAGCATT 416  
QY 461 TTAGGTCAAGCAGACGACAAAGAGAGATTAAGTGGCGCACTCCCGGTGGCATCACTTGCAT 520  
Db 417 TTAGGTCAAGCAGACGACAAAGAGAGATTAAGTGGCGCACTCCCGGTGGCATCACTTGCAT 476  
QY 521 CTTGAGAGACGACCTTGTGTAAGTCAAGAGCTCTACACATTGAATTCGAAGTCAAGCTTC 580

Db 477 CCTTGGAGCAGACCTTGTGAAGTCAAGAGCTCTACACATTGAATTCGAAGTCAAGCTTC 536  
QY 581 TGAATTCGATGATTCAGAGAGAAAGAGATTCGAGTTTGAAGCAAAATGAGTGAGACA 640  
Db 537 TGAATTCGATGATTCAGAGAGAAAGAGATTCGAGTTTGAAGCAAAATGAGTGAGACA 596  
QY 641 GCACCAATTCGAATGTTGTAATTTGAAGACTTGAAGTCTCACTTACCCGTGGTACCCC 700  
Db 597 GCACCAATTCGAATGTTGTAATTTGAAGACTTGAAGTCTCACTTACCCGTGGTACCCC 656  
QY 701 CACTTCACAAGCTCTCTGAATATATCTGTTTACCCCGAGAAATATTACAGAGTATC 760  
Db 657 CACTTCACAAGCTCTCTGAATATATCTGTTTACCCCGAGAAATATTACAGAGTATC 716  
QY 761 TGCAGAGTCTCTTGGCCAGG 781  
Db 717 TGCAGAGTCTCTTGGCCAGG 737

RESULT 11  
AC010346 109626 bp DNA linear PRI 10-NOV-2000  
LOCUS AC010346 Homo sapiens chromosome 5 clone CITB-H1\_2018B2, complete sequence.  
ACCESSION AC010346  
VERSION AC010346.6 GI:11136705  
KEYWORDS HTG.  
SEGMENT Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 109626)  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 109626)  
AUTHORS DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
3 (bases 1 to 109626)  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Submitted (10-NOV-2000) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
On Nov 10, 2000 this sequence version replaced gi:9256196.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www-shgc.stanford.edu  
Quality: Phrap Quality >=40 99.9% of Sequence;  
STS Content:  
WI-9255 G05749.

FEATURES  
source  
1. .109626  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/clone="CITB-H1\_2018B2"

ORIGIN  
Query Match 15.8%; Score 330; DB 9; Length 109626;  
Best Local Similarity 99.7%; Pred. No. 2.2e-168;  
Matches 380; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 401 GTTTAGAACTTGTGTGAGCCGTGATTTGCTGAGCTCTGGCCAGCCCTCAGAAAGCATT 460  
Db 88571 GTTTAGAACTTGTGTGAGCCGTGATTTGCTGAGCTCTGGCCAGCCCTCAGAAAGCATT 88630  
QY 461 TTAGGTCAAGCAGACGACAAAGAGAGATTAAGTGGCGCACTCCCGGTGGCATCACTTGCAT 520  
Db 88631 TTAGGTCAAGCAGACGACAAAGAGAGATTAAGTGGCGCACTCCCGGTGGCATCACTTGCAT 88690

Qy	52	CCTTGAAGACAGACCTTGTGAAGTCAAGCTGCATCATTTGAATCTCAAGTGAAGCTTC	580
Db	88691	CCTGAGAGACAGACCTTGTGAAGTCAAGCTGCATCATTTGAATCTCAAGTGAAGCTTC	88755
Qy	581	TGAGATTCGATGATTCAGAGAAAGGATTCAGAGTTTGAAGCAAAATGCAAGTGAACA	640
Db	88751	TGAGATTCGATGATTCAGAGAAAGGATTCAGAGTTTGAAGCAAAATGCAAGTGAACA	88810
Qy	641	GCAACCAATCAATGTTGTAATGAACAATTGAGTCTACATTAACCGTTCGGTACCCC	700
Db	88811	GCAACCAATCAATGTTGTAATGAACAATTGAGTCTACATTAACCGTTCGGTACCCC	88870
Qy	701	CACCTCAACAAGCCTCTCTGAATATTCCTGATTTACCCCCAGAAATATTTACAGTACATC	760
Db	88871	CACCTCAACAAGCCTCTCTGAATATTCCTGATTTACCCCCAGAAATATTTACAGTACATC	88930
Qy	761	TGCAGAGATCTCTTGGCCAGG	781
Db	88931	TGCAGAGATCTCTTGGCCAGG	88951

RESULT	12
AC025174	
LOCUS	AC025174
DEFINITION	AC025174 Homo sapiens chromosome 5 clone CTD-207I24, complete sequence.
ACCESSION	AC025174
VERSION	AC025174.5
KEYWORDS	GI:19774456
SOURCE	Htg.
ORGANISM	Homo sapiens (human)
	Homo sapiens

REFERENCE	1 (bases 1 to 110756)
AUTHORS	DOE Joint Genome Institute and Stanford Human Genome Center
TITLE	Direct Submission

**JOURNAL** Unpublished  
**REFERENCE** 2 (bases 1 to 110756)

**AUTHORS** DOE Joint Genome Institute.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (07-MAR-2000) Production Sequencing Facility, DOE Joint

**REFERENCE**  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
**AUTHORS**  
3 (bases 1 to 110756)  
DOJ Joint Genome Institute.

**TITLE** Direct Submission  
**SUBMITTED** (07-MAR-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
**JOURNAL**

REFERENCE  
DOE Joint Genome Institute and Stanford Human Genome Center.  
4 (bases 1 to 110756)  
AUTHORS  
TITLE  
Direct Submission  
JOURNAL  
Submitted (28-MAR-2002) DOE Joint Genome Institute, 2800 Mitchell

COMMENT  
On Mar 28, 2002 this sequence replaced gi:19224767.  
Drive, Walnut Creek, CA 94598, USA

draft sequence produced by DOE Joint Genome  
[www.jgi.doe.gov](http://www.jgi.doe.gov)  
 Finishing Completed at Stanford Human Genome  
[www.sngc.stanford.edu](http://www.sngc.stanford.edu)  
 Quality: Phrap Quality >=40 100% of Sequence  
 Estimated Total Number of Errors is 0.

FEATURES	Location/Qualifiers
source	1. .110756

```

/organism="Homo sapiens"
/mol_type="Genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2072124"

```

Query Match 15.8%; Score 330; DB 9; Length 110756;  
 Bect Local Similarity 99.7%; Pred. No. 2.2e-168;  
 Matches 380; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 401 GTTTGAACTTGTGTTGAGCCCTGAGATTGCTGGAGACTGCCACCTCTGAGAAAGCATTT 460

Db	20040	GTTTAGAACTTGGTGTGAGCCGTGATTTGCTGACCTTGGCCAGACGCTCAGAAAGCATT	200393
QY	461	TTAGGTCAAGCAGAGGACCAAGAGGAATTAAGTGGCCCATCTCCCGGTGGCATCACTTCGCAT	520
Db	20100	TTAGGTCAAGCAGAGGACCAAGAGGAATTAAGTGGCCCATCTCCCGGTGGCATCACTTCGCAT	201558
QY	521	CTTTGAGGAGAGACCTTTGGAAGTCAAGCTGCTACATTTGAATCTCAAGTCAAGCTTC	580
Db	20160	CTTGGAGAGAGACCTTTGGAAGTCAAGCTGCTACATTTGAATCTCAAGTCAAGCTTC	202191
QY	581	TGAGATTGATGATTCAGGAAAGAAAGATTTCTAGGTTTGAAGCAAAATGCAGTGAACA	640
Db	20220	TGAGATTGATGATTCAGGAAAGAAAGATTTCTAGGTTTGAAGCAAAATGCAGTGAACA	202778
QY	641	GCAACCAATCCAATGTTGTATTTGAAGACTTTGAGTCTCACTTACCCTGGTACCC	700
Db	20280	GCAACCAATCCAATGTTGTATTTGAAGACTTTGAGTCTCACTTACCCTGGTACCC	203380
QY	701	CACCTTCACAAGCCTCTCTGAAATATTCCTGCTTACCCCCAGAAATATTTACAGATATC	760
Db	20340	CACCTTCACAAGCCTCTCTGAAATATTCCTGCTTACCCCCAGAAATATTTACAGATATC	203995
QY	761	TGCAGAGATCTCTTGGCCAGG	781
Db	20400	TGCAGAGATCTCTTGGCCAGG	204420

LOCUS	DEFINITION	AC022921	158199 bp	DNA	linear	HTG 12-MAR-2000
AC022921	Homo sapiens clone RP11-138P20, WORKING DRAFT SEQUENCE, 12 unordered pieces.					

ACCESSION	AC022921
VERSION	AC022921.2
KEYWORDS	GI:7229868
	HTG; HTGS_PHASE1; HTGS_DRAFT

**SOURCE**  
**ORGANISM**

REFERENCE  
1 (bases 1 to 158199)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

unpublished  
Homo sapiens, clone RP11-138P20  
and Lander, E.  
Nusbaum, C.  
Linton, J.,  
Britten, B.

REFERENCE  
AUTHORS

2 (bases 1 to 158,199)  
Barren, B., Linton, J., Nusbaum, C., Lander, E., Abraham, H., Allen, N.  
Anderson, S., Baldwin, J., Barna, N., Beckert, R., Beda, F.,

Bogunavsky, L., Boukngatei, B., Brown, A., Burrell, C., Caselle, W., Choepl, Y., Collangelo, M., Collins, S., Collamore, A., Cooke, P., DeRrellano, K., Dewar, K., Domingo, M., Doyle, M., Fenevor, J., Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J.,

Gardyna, S., Grant, G., Hagos, B., Heatford, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kann, L., Karates, A., Klein, J.

Mander, B., Denochy, O., Levine, K., Ueda, C., Ding, S., Locke, K., MacDonald, P., Margulis, N., McWan, P., McGuck, A., McKernan, K., McHeesters, R., Meldrim, J., Meneis, L., Morrow, J., Naylor, J., Notman, C. H., O'Connor, T., O'Donnell, P., Oliver, T. M., Peterson, K., Pierre, N., Pisanti, C., Pollara, V., Raymond, C., Riley, R., Rothman, L., Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,

Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W., Zeng, J., Stojanovic, N., Sudramanah, A., Talamas, C., Terstegen, S., Theodorou, C., Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W., Zeng, J.

TITLE	Zimmer, A. and Zody, M.
JOURNAL	Direct Submission
COMMENT	Submitted (07-FEB-2000) Whitehead Institute/MIT Center for Genomomms Research, 320 Charles Street, Cambridge, MA 02141, USA On Mar 12, 2000 this sequence version replaced g1:6921809. All repeats were identified using RepeatMasker:

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
 ----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: <http://www-seq.wi.mit.edu>



misc_feature	/note="assembly_fragment"
36860..53921	/note="assembly_fragment"
misc_feature	/note="assembly_fragment"
54022..72054	/note="assembly_fragment"
misc_feature	/note="assembly_fragment"
72155..102527	/note="assembly_fragment"
misc_feature	/note="assembly_fragment"
102628..158199	/note="assembly_fragment"
ORIGIN	
Query Match	13.3%; Score 279; DB 2; Length 158199;
Best Local Similarity	99.5%; Pred. No. 1,7e-140;
Matches 379; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
QY	401 GTTTAGAACTTGTGGTTGAGCCCGTGAATTCGTGACTCTGCGCAGCCCTCAGAAAGCATT 460
Db	3446 GTTTAGAACTTGTGGTTGAGCCCGTGAATTCGTGACTCTGCGCAGCCCTCAGAAAGCATT 3505
QY	461 TTAGGTCAACGACAGGACAGAGAGAGTAGTGGCGGACCTCCCGGTGGCATTCACCTGGCAT 520
Db	3506 TTAGGTCAACGACAGGACAGAGAGAGTAGTGGCGGACCTCCCGGTGGCATTCACCTGGCAT 3565
QY	521 CCTTGAGACAGACCTTGTGAAGTCAGAGCGCTGACATTTGAATCTCAAGTCGAGCTTC 580
Db	3566 CCTTGAGACAGACCTTGTGAAGTCAGAGCGCTGACATTTGAATCTCAAGTCGAGCTTC 3625
QY	581 TGAGATTGATGATTCAGAGAAAGGATTCGAGTTTGAAGCAAAATGCAATGAGTGAACA 640
Db	3626 TGAGATTGATGATTCAGAGAAAGGATTCGAGTTTGAAGCAAAATGCAATGAGTGAACA 3685
QY	641 GCAACCAATCAATGTTGTAAATGGAAGACTTTGAGTCTCACTTACCCGTTGGTACCCC 700
Db	3686 GCAACCAATCAATGTTGTAAATGGAAGACTTTGAGTCTCACTTACCCGTTGGTACCCC 3745
QY	701 CACTCTCAAGAGCCCTCTGAATATTCCTGAGTTTACCCCGAATATTTACAGTACATC 760
Db	3746 CACTCTCAAGAGCCCTCTGAGATATTCCTGAGTTTACCCCGAATATTTACAGTACATC 3805
QY	761 TGCAGGAGTCTCTTGGCCAGG 781
Db	3806 TGCAGGAGTCTCTTGGCCAGG 3826
RESULT 14	
AC021609	167237 bp DNA linear HTG 12-MAR-2000
LOCUS	
DEFINITION	Homo sapiens clone RP11-259D10, WORKING DRAFT SEQUENCE, 6 unordered pieces.
ACCESSION	AC021609
VERSION	AC021609.3 GI:7230210
SOURCE	HTG; HTGS PHASE1; HTGS_DRAFT.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE	1 (bases 1 to 167237)
JOURNAL	Bitren,B., Linton,L., Nusbaum,C. and Lander,E.
REFERENCE	Unpublished
AUTHORS	2 (bases 1 to 167237)
	Bitren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,D., Barna,N., Beckertly,R., Beda,F., Boguslavsky,L., Bouhgalter,B., Brown,A., Burkett,G., Castle,A., Choepel,J., Collangelo,M., Collins,S., Collymore,A., Cooke,P., Dekrellano,K., Dewar,K., Domino,M., Doyle,M., Feneceor,J., Ferreira,S., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J., Gadyra,S., Grant,G., Hagos,B., Heatford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Margulis,N., McEwan,P., McGuck,A., McKernan,K., McHesters,R., Meldrim,T., Menneus,L., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K.,





Quality coverage: 10.01 in Q20 bases; sum-of-contigs estimation.

\* NOTE: This is a 'working draft' sequence. It currently consists of 27 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. \* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1207: contig of 1207 bp in length  
1308: gap of unknown length  
1308: contig of 2078 bp in length  
3386: gap of unknown length  
3486: contig of 2068 bp in length  
5553: gap of unknown length  
5654: contig of 1377 bp in length  
7030: gap of unknown length  
7130: contig of 1939 bp in length  
9069: gap of unknown length  
9169: contig of 2283 bp in length  
9170: gap of unknown length  
11452: contig of 4450 bp in length  
11553: gap of unknown length  
16002: contig of 4981 bp in length  
16103: gap of unknown length  
21083: contig of 4981 bp in length  
21084: gap of unknown length  
21183: contig of 5710 bp in length  
26893: gap of unknown length  
26993: contig of 2994 bp in length  
29987: gap of unknown length  
30088: contig of 2862 bp in length  
32950: gap of unknown length  
33050: contig of 5708 bp in length  
38578: gap of unknown length  
38858: contig of 6345 bp in length  
45202: gap of unknown length  
45302: contig of 5609 bp in length  
45303: gap of unknown length  
50912: contig of 4869 bp in length  
51012: gap of unknown length  
55880: contig of 7209 bp in length  
55981: gap of unknown length  
63189: contig of 6574 bp in length  
63290: gap of unknown length  
63290: contig of 6343 bp in length  
69633: gap of unknown length  
69732: contig of 6645 bp in length  
76378: gap of unknown length  
76378: contig of 6574 bp in length  
83052: gap of unknown length  
83152: contig of 11778 bp in length  
94930: gap of unknown length  
95029: contig of 9317 bp in length  
104346: gap of unknown length  
104447: contig of 10829 bp in length  
115275: gap of unknown length  
115376: contig of 8113 bp in length  
123489: gap of unknown length  
123589: contig of 9584 bp in length  
133172: gap of unknown length  
133273: contig of 12223 bp in length  
145495: gap of unknown length  
145596: contig of 14675 bp in length  
160271: gap of unknown length  
160371: contig of 17226 bp in length.

FEATURES

source  
1.177596  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/clone\_lib="RP11-35616"  
/clone\_lib="RP11-35616" human BAC library 11"

ORIGIN

Query Match 13.3%; Score 279; DB 2; Length 177596;

Best Local Similarity 99.5%; Pred. No. 1.7e-140; Matches 379; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 401 GTTTAGAACCTTGTTGGTGGCCGTGATTTGCTGACCTGGCCAGCCCTCAGAAACATT 460  
Db 83524 GTTTAGAACCTTGTTGGTGGCCGTGATTTGCTGACCTGGCCAGCCCTCAGAAACATT 83583  
QY 461 TTAGGTCAAGCAGAGGACAAAGAGATTAAGTGGCCACTCCCGTGGCATCCTGCAT 520  
Db 83584 TTAGGTCAAGCAGAGGACAAAGAGATTAAGTGGCCACTCCCGTGGCATCCTGCAT 83643  
QY 521 CCTTGAGACAGACCTTGTGAAGTCAAGCTGTACACATTAATTCATAGTGAAGTTC 580  
Db 83644 CCTTGAGACAGACCTTGTGAAGTCAAGCTGTACACATTAATTCATAGTGAAGTTC 83703  
QY 581 TGAGATTGATGATTCAGGAGAAAGATTTGAGTTTGAAGCAAAATGACGTGACA 640  
Db 83704 TGAGATTGATGATTCAGGAGAAAGATTTGAGTTTGAAGCAAAATGACGTGACA 83763  
QY 641 GCACCAATCCAAATGTTGTAATTGAAGACTTGAATCCTCACTTACCCGTTGGTACCC 700  
Db 83764 GCACCAATCCAAATGTTGTAATTGAAGACTTGAATCCTCACTTACCCGTTGGTACCC 83823  
QY 701 CACTTCACAGGCTCTCTGAATATTCCTGTTTACCCCAAGATTTTACAGTATC 760  
Db 83824 CACTTCACAGGCTCTCTGAATATTCCTGTTTACCCCAAGATTTTACAGTATC 83883  
QY 761 TGCAGAGTCTCTGSCCAGG 781  
Db 83884 TGCAGAGTCTCTGSCCAGG 83904

Search completed: August 27, 2005, 09:39:04  
Job time : 5999.75 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 13:32:20 ; Search time 732.091 Seconds  
(without alignments)  
16924.161 Million cell updates/sec

Title: US-09-371-347A-47

Perfect score: 2093

Sequence: 1 atgagggaggttcctgtact.....ttcagatattgtgcataa 2093

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq.16Dec04:\*

1: geneeqn1980s:\*\n2: geneeqn1990s:\*\n3: geneeqn2000s:\*\n4: geneeqn2001as:\*\n5: geneeqn2001bs:\*\n6: geneeqn2002as:\*\n7: geneeqn2002bs:\*\n8: geneeqn2003as:\*\n9: geneeqn2003bs:\*\n10: geneeqn2003cs:\*\n11: geneeqn2003ds:\*\n12: geneeqn2004as:\*\n13: geneeqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1970	94.1	2091	11	ADM43214 Human met
2	1793	85.7	3259	5	AA65070 DNA encod
3	1793	85.7	3259	5	AA65070 DNA encod
4	1793	85.7	3259	11	ADM43206 Human scd
5	1793	85.5	2094	11	ADM43208 Human wll
6	1739	83.1	2094	11	ADM43212 Human met
7	1739	83.1	2094	11	ADM43209 Human met
8	1691	80.8	3259	3	AA58935 DNA encod
9	1677	80.1	2091	11	ADM43216 Human met
10	1640	78.4	3270	13	ADQ87538 Human tum
11	1595	76.2	3256	3	AA58977 A human m
12	1544	73.8	3189	13	ACN42470 Human dia
13	956	45.7	3256	13	ADQ39029 Human SNP
14	879	42.0	3274	13	ADQ39030 Human SNP
15	879	42.0	3274	13	ADQ39030 Human SNP
16	501	23.9	1986	4	AA541064 DNA enco
17	386	18.4	390	2	AA51820 Human sec
18	330	15.7	591	12	ACH73174 Human Gen
19	328	15.7	379	12	ACH86905 Human Gen
20	317	15.1	1663	4	AA541602 cDNA enco

21	279	13.3	591	12	ACH68540 Human gen
22	277	13.2	379	12	ACH82240 Human gen
23	225	10.8	503	5	AA65069 DNA encod
24	188	9.0	525	12	ACH67438 Human gen
25	175	8.4	175	12	ACH81143 Human gen
26	158	7.5	2475	6	AD32365 Human lun
27	157	7.5	2475	13	AD161720 Human gen
28	137	6.5	525	12	ACH73117 Human gen
29	124	5.7	175	12	ACH86848 Human gen
30	78	3.9	244	3	AA242736 Human 5'
31	60	2.9	60	6	AB36264 Human sp1
32	51	2.4	51	4	AA178548 Human s11
33	38	1.8	1835	5	AA65071 DNA encod
34	30	1.4	1681	11	AD131127 Human CDN
35	26	1.2	26	3	AA58955 PCR prime
36	26	1.2	26	3	AA58939 PCR prime
37	26	1.2	26	6	ABX09549 Arteriosc
38	26	1.2	26	6	AA143713 Pregelstat
39	26	1.2	26	11	ADM43205 Human met
40	26	1.2	26	11	ADM43189 Human met
41	25	1.2	25	3	AA58952 PCR prime
42	25	1.2	25	3	AA58937 PCR prime
43	25	1.2	25	3	AA58947 PCR prime
44	25	1.2	25	11	ADM43187 Human met
45	25	1.2	25	11	ADM43202 Human met

## ALIGNMENTS

RESULT 1	ADM43214	standard; cDNA; 2091 BP.
ID	ADM43214	standard; cDNA; 2091 BP.
XX	ADM43214;	
AC	03-JUN-2004	(first entry)
DT	03-JUN-2004	(first entry)
XX	Human methionine synthase reductase CDS del 1675-1678 variant.	
DB	Human; s8; Methionine synthase reductase polypeptide; HAMTRR; cancer;	
KW	cardiovascular disease; neural tube defect; hyperhomocysteinemia;	
KW	chromosome 5p15.2-p15.3; SNP; single nucleotide polymorphism.	
XX	Homo sapiens.	
OS	Homo sapiens.	
XX	Key	Location/Qualifiers
FT	CDS	1..2091
FT		/*tag= a
FT		/product= "HAMTRRdelR559"
FT		/partial
FT		/note= "No stop codon shown"
FT	variation	replace(66, A)
FT		/*tag= b
FT		/standard name= "single_nucleotide polymorphism"
FT	variation	replace(110, A)
FT		/*tag= c
FT		/standard name= "single_nucleotide polymorphism"
FT	variation	replace(1675, AGAG)
FT		/*tag= d
PN	US2003082676-A1.	
XX	01-MAY-2003.	
PD	10-AUG-1999;	99US-00371347.
XX	15-JAN-1998;	98US-0071622P.
PR	15-JAN-1999;	99US-00232028.
XX	(GRAV/) GRAVEL R A.	
PA	(ROZE/) ROZEN R.	
PA	(LECL/) LECLERC D.	



Qy	1561	CCATCTCTCGTGAACAACAATTTCTTTCACATTAACAATGACCCCTCAATCCCATC	1620
Db	1561	TCATCTCTCTCGTGAACAACAATTTCTTTCACATTAACAATGACCCCTCAATCCCATC	1620
Qy	1621	ATAATGCTGGTCTCCAGAACCCGGCATAGCCCGCTTATTTGGGTTCTTAACAAT-AGAA	1679
Db	1621	ATAATGCTGGTCTCCAGAACCCGGCATAGCCCGCTTATTTGGGTTCTTAACAATAGAAA	1680
Qy	1680	CTCCAGAACAAACCCAGATGAAATTTTGGAGCAATGTGGTGTCTTTTGGCTGCAGG	1739
Db	1681	CTCCAGAACAAACCCAGATGAAATTTTGGAGCAATGTGGTGTCTTTTGGCTGCAGG	1740
Qy	1740	CATAAGATAGGAGATTATCTATTGAGAAAAGAGCTCAGACATTTCTTAAGATGGAGTC	1799
Db	1741	CATAAGATAGGAGATTATCTATTGAGAAAAGAGCTCAGACATTTCTTAAGATGGAGTC	1800
Qy	1800	TTAACTCATCTAAAGGTTTCTCTTCCTTCAAGAGATGCTCCTGTGGGAGAGAAAGCCCA	1855
Db	1801	TTAACTCATCTAAAGGTTTCTCTTCCTTCAAGAGATGCTCCTGTGGGAGAGAAAGCCCA	1860
Qy	1860	GCAAAAGTATGTAACAAGACAATCTCAGTTCAATGGCCACAGGTGGCGAAGATCCTCTC	1919
Db	1861	GCAAAAGTATGTAACAAGACAATCTCAGTTCAATGGCCACAGGTGGCGAAGATCCTCTC	1920
Qy	1920	CAGAGAAACGGCCATATTTATGTGTGTGGAGTGCAGAAAGAAATATGGCCAAAGATACAT	1979
Db	1921	CAGAGAAACGGCCATATTTATGTGTGTGGAGTGCAGAAAGAAATATGGCCAAAGATACAT	1980
Qy	1980	GATGCCCTTGTGTCAAAATATATAAGCAAAAGGTTGGAGTTGAAAACTAGAACATGAAA	2039
Db	1981	GATGCCCTTGTGTCAAAATATATAAGCAAAAGGTTGGAGTTGAAAACTAGAACATGAAA	2040
Qy	2040	ACCTGTGCTACTTTTAAAGAAAGAAAAGCTACCTTCAGATATTTGGTCA	2090
Db	2041	ACCTGTGCTACTTTTAAAGAAAGAAAAGCTACCTTCAGATATTTGGTCA	2091
RESULT 2			
AA	AA565070		
ID	AA565070	standard; cDNA; 3259 BP.	
XX	AA565070;		
AC	AA565070;		
XX	13-FEB-2002	(first entry)	
DT	13-FEB-2002	(first entry)	
XX	DNA encoding novel human diagnostic protein #874.		
DE	DNA encoding novel human diagnostic protein #874.		
XX	Human; chromosome mapping; gene mapping; gene therapy; forensic;		
KV	Human; chromosome mapping; gene mapping; gene therapy; forensic;		
KM	food supplement; medical imaging; diagnostic; genetic disorder; ss.		
XX	Homo sapiens.		
OS	Homo sapiens.		
XX	WO200175067-A2.		
FN	WO200175067-A2.		
PD	11-OCT-2001.		
XX	11-OCT-2001.		
PF	30-MAR-2001; 2001WO-US008631.		
XX	30-MAR-2001; 2001WO-US008631.		
XX	31-MAR-2000; 2000US-00540217.		
PR	23-AUG-2000; 2000US-00649167.		
XX	(HYSB-) HYSBQ INC.		
PA	(HYSB-) HYSBQ INC.		
XX	Dzmanac RT, Liu C, Tang YT;		
PI	Dzmanac RT, Liu C, Tang YT;		
DR	WPI; 2001-639362/73.		
DR	P-PSDB; ABG00883.		
XX	New isolated polynucleotide and encoded polypeptides, useful in		
PT	diagnostics, forensics, gene mapping, identification of mutations		
PT	responsible for genetic disorders or other traits and to assess		
PT	biodiversity.		

Query Match	Beat Local Similarity	85.7%; Score 1793;	DB 5;	Length 3259;				
Matches 2993;	Conservative	0;	Mismatches	0;	Indels	4;	Gaps	1.
1	ATGAGAGAGTTCTGTGTAATGCTATATGCTACACAGCAGGAGGACGCAAGGCCATCGCAGAA	60						
80	ATGAGAGAGTTCTGTGTAATGCTATATGCTACACAGCAGGAGGACGCAAGGCCATCGCAGAA	139						
61	GAATGCTGAGCAGCTGTGTATCATGATTTCTGAGATCTTCACTGTATTAGTGA	120						
140	GAATGCTGAGCAGCTGTGTATCATGATTTCTGAGATCTTCACTGTATTAGTGA	199						
121	TCGGATAAGTATGACTTAATAAACCGAAACAGCTCTCTGTGTGTGTGTTCTACACAG	180						
200	TCGGATAAGTATGACTTAATAAACCGAAACAGCTCTCTGTGTGTGTGTTCTACACAG	259						
181	GCGACCGGAGACCCACCCGACACAGCCCGCAAGTTGTTAAGAAATACAGACCAACA	240						
260	GCGACCGGAGACCCACCCGACACAGCCCGCAAGTTGTTAAGAAATACAGACCAACA	319						
241	CTGCCGCTGTAATTTCTTGTCTACCTGGGGTATGGGTTACCTGGGCTCGATTCAGAA	300						
320	CTGCCGCTGTAATTTCTTGTCTACCTGGGGTATGGGTTACCTGGGCTCGATTCAGAA	379						
301	TACACCTTCTTTTGCAATGGGGGGGAAATTAATGATTAACGACTTCAAGAGCTTGGAGCC	360						
380	TACACCTTCTTTTGCAATGGGGGGGAAATTAATGATTAACGACTTCAAGAGCTTGGAGCC	439						
361	CGGCAATTTCTATGACACTGACATGACAGATGACTGTGTAGGTTTGAACCTGTGGTTAG	420						
440	CGGCAATTTCTATGACACTGACATGACAGATGACTGTGTAGGTTTGAACCTGTGGTTAG	499						
421	CGGTGGATTGCTGACTCTGGCCAGCCCTCAGAAAGCAATTTTAGTCAAGCAGAGACAA	480						
500	CGGTGGATTGCTGACTCTGGCCAGCCCTCAGAAAGCAATTTTAGTCAAGCAGAGACAA	559						
481	GAGGAGATTAAGGGGCACTCCGGGTGACATCACTGATCCTTTAGAGACAGACTTGTG	540						
560	GAGGAGATTAAGGGGCACTCCGGGTGACATCACTGATCCTTTAGAGACAGACTTGTG	619						
541	AAGTCAGAGCTGCTACACATTGAATCTCAAGTCGAGCTTCTGAGATTCGATGATTCAGAA	600						
620	AAGTCAGAGCTGCTACACATTGAATCTCAAGTCGAGCTTCTGAGATTCGATGATTCAGAA	679						
601	AGAAAGATTTGAGGTTTGAAGCAAAATGCAAGTGAACAGAACCAATCCAAATGTTGTA	660						

Db 680 AGAAGATTCTGAGTTTGAAGCAAAATGACGTGAACAAACCAATCCAAATGTTGA 739  
 Qy 661 ATTGAAGACTTTGAGTCTGACTTACCCGTGGTACCCCACTCTCAAGGCTCTG 720  
 Db 740 ATTGAAGACTTTGAGTCTGACTTACCCGTGGTACCCCACTCTCAAGGCTCTG 739  
 Qy 721 AATATTCTGTTTACCCCAAGAAATTTACAGGTACATGCGAGAGTCTTTGGCAG 780  
 Db 800 AATATTCTGTTTACCCCAAGAAATTTACAGGTACATGCGAGAGTCTTTGGCAG 859  
 Qy 781 GAGGAAGCCAAAGTCTGAGCTTGAAGCAATGCTTTTGAAGGCCAATTTCAAG 840  
 Db 860 GAGGAAGCCAAAGTCTGAGCTTGAAGCAATGCTTTTGAAGGCCAATTTCAAG 919  
 Qy 841 GCAATTCACCTTACTAGAAATGATGCAATTAACCACTGCTGTGTAATTTGAAT 900  
 Db 920 GCAATTCACCTTACTAGAAATGATGCAATTAACCACTGCTGTGTAATTTGAAT 979  
 Qy 901 TCAAAATACAGCTTTTCTATCAGCTGAGATGCTTCAAGCTGATCTGCTAACAT 960  
 Db 980 TCAAAATACAGCTTTTCTATCAGCTGAGATGCTTCAAGCTGATCTGCTAACAT 1039  
 Qy 961 GATTCTGAGGTACAAAGCTACTCCAAAGCTGAGCTTGAAGTAAGAGAGCACTGC 1020  
 Db 1040 GATTCTGAGGTACAAAGCTACTCCAAAGCTGAGCTTGAAGTAAGAGAGCACTGC 1099  
 Qy 1021 GTCTTTTGAATAAAGGAGACACAAAGAAAGAGACTACTTACCCAGCAATATA 1080  
 Db 1100 GTCTTTTGAATAAAGGAGACACAAAGAAAGAGACTACTTACCCAGCAATATA 1159  
 Qy 1081 CTGCGGGAGTGTCTCTCAAGTTTATCTGCTGCTTGAATCCAGCAATTCCT 1140  
 Db 1160 CTGCGGGAGTGTCTCTCAAGTTTATCTGCTGCTTGAATCCAGCAATTCCT 1219  
 Qy 1141 AAAAAGGATTTTGGAGGCTTGGAGCTATACAGTACAGTGCTGAAAAGCGCAGG 1200  
 Db 1220 AAAAAGGATTTTGGAGGCTTGGAGCTATACAGTACAGTGCTGAAAAGCGCAGG 1279  
 Qy 1201 CTACAGAGCTGTGCACTAAACAAAGGAGCGATTAAGCGCTTTGTAACGAGATGCC 1260  
 Db 1280 CTACAGAGCTGTGCACTAAACAAAGGAGCGATTAAGCGCTTTGTAACGAGATGCC 1339  
 Qy 1261 TGTGCTGCTTTGGATCTCTCTCTGCTTCCCTTCCAGCCACCACTCAGTCTC 1320  
 Db 1340 TGTGCTGCTTTGGATCTCTCTCTGCTTCCCTTCCAGCCACCACTCAGTCTC 1399  
 Qy 1321 CTGCTCGAATCTTCTCTAACTCAACCCAGACATATTGATGCAAGCTCAAGTTTA 1380  
 Db 1400 CTGCTCGAATCTTCTCTAACTCAACCCAGACATATTGATGCAAGCTCAAGTTTA 1459  
 Qy 1381 TTTCACCCAGAAAAGCTCAATTTGTCTTCAACATTTGTGTAATTTCTCTACGCCACA 1440  
 Db 1460 TTTCACCCAGAAAAGCTCAATTTGTCTTCAACATTTGTGTAATTTCTCTACGCCACA 1519  
 Qy 1441 ACAGAGTGTCTGCGGAAGGAGATATGTAACAGCTGCTGCTGCTTTGTTGCTTCAAGTT 1500  
 Db 1520 ACAGAGTGTCTGCGGAAGGAGATATGTAACAGCTGCTGCTGCTTTGTTGCTTCAAGTT 1579  
 Qy 1501 CTTCAGCAAAATACATGATCCATGAGAGACAGCGGAAAGCCGTGCTCCTAAGATA 1560  
 Db 1580 CTTCAGCAAAATACATGATCCATGAGAGACAGCGGAAAGCCGTGCTCCTAAGATA 1639  
 Qy 1561 TCCATCTCTCTCGAACAACAAATTTCTTCCACTTACAGATGACCCCTCAATCCCATC 1620  
 Db 1640 TCCATCTCTCTCGAACAACAAATTTCTTCCACTTACAGATGACCCCTCAATCCCATC 1699  
 Qy 1621 ATTAATGTTGGTCCAGAACCGGATATGCCCCGTTTATTTGGGTTCTTACAAAT---AG 1676  
 Db 1700 ATTAATGTTGGTCCAGAACCGGATATGCCCCGTTTATTTGGGTTCTTACAAATAGAGAG 1759  
 Qy 1677 AAATCTCAAGAACAAACCCAGATGAGAAATTTTGAACCAATGAGTGTGTTTGGCTGC 1736  
 Db 1760 AAATCTCAAGAACAAACCCAGATGAGAAATTTTGAACCAATGAGTGTGTTTGGCTGC 1819

Qy 1737 AGGCAATAGAGATAGGATTAATCTATTCAAGAAAAGACTCAGACATTTCTTAAGCATGG 1796  
 Db 1820 AGGCAATAGAGATAGGATTAATCTATTCAAGAAAAGACTCAGACATTTCTTAAGCATGG 1879  
 Qy 1797 ATCTTAACCTATCTAAAGGTTTCTTCTCAAGAGATGCTCTGTGGGAGAGAGAAAGCC 1856  
 Db 1880 ATCTTAACCTATCTAAAGGTTTCTTCTCAAGAGATGCTCTGTGGGAGAGAGAAAGCC 1939  
 Qy 1857 CCAGCAAGATATGTAACAACAACATCCAGCTTCATGGCCAGCAGGTGGCGAATCTCTC 1916  
 Db 1940 CCAGCAAGATATGTAACAACAACATCCAGCTTCATGGCCAGCAGGTGGCGAATCTCTC 1999  
 Qy 1917 CTTCAGAGAACGCGCATATTTATGTGTGAGATGCAAGAAATATGCGCAAGATGTA 1976  
 Db 2000 CTTCAGAGAACGCGCATATTTATGTGTGAGATGCAAGAAATATGCGCAAGATGTA 2059  
 Qy 1977 CATGATGCCCCCTGTGCAATTAATTAACCAAGAGTTGGAGTTGAAAACCTAGAACCAATG 2036  
 Db 2060 CATGATGCCCCCTGTGCAATTAATTAACCAAGAGTTGGAGTTGAAAACCTAGAACCAATG 2119  
 Qy 2037 AAAACCTGGCCACTTTAAAGAAAGAAACGCTACCTTCAGAGATATTTGTCATTA 2093  
 Db 2120 AAAACCTGGCCACTTTAAAGAAAGAAACGCTACCTTCAGAGATATTTGTCATTA 2176

RESULT 3  
 ID AAC91226  
 AAC91226 standard; DNA; 3259 BP.  
 AAC91226;  
 20-MAR-2001 (first entry)  
 DE Human schizophrenia related gene SHQ ID NO: 23.  
 KM Human; schizophrenia; developmental disorder; spina bifida cystica;  
 KM Tourette's syndrome; bipolar illness; autism; conduct disorder;  
 KM attention deficit disorder; obsessive compulsive disorder;  
 KM chronic multiple tic syndrome; learning disorder; polymorphism; ds.  
 OS Homo sapiens.  
 PN W0200071754-A1.  
 30-NOV-2000.  
 PF 24-MAY-2000; 2000MC-US014354.  
 PR 25-MAY-1999; 99US-00318448.  
 PA (UNB-) UNIV NEW JERSEY MEDICINE & DENTISTRY.  
 PI Johnson WG, Stenroos ES;  
 DR MPI; 2001-025174/03.  
 XX  
 PT Diagnosing a developmental disorder, e.g. schizophrenia, by forming  
 PT datasets (DS) of genetic (e.g. genotypes of folate metabolism alleles)  
 PT and environmental variables affecting an individual and then comparing  
 PT these DS with reference DS.  
 XX  
 PS Disclosure; Page 142-143; 156pp; English.  
 CC The present invention provides a novel method of estimating the  
 CC susceptibility of an individual to a developmental disorder using genetic  
 CC and environmental variables. The method can be used in the diagnosis,  
 CC prevention and treatment of disorders such as schizophrenia, spina bifida  
 CC cystica, Tourette's syndrome, bipolar illness, autism, conduct disorders,  
 CC attention deficit disorder, obsessive compulsive disorder, chronic  
 CC multiple tic syndrome and learning disorders such as dyslexia  
 CC  
 SO Sequence 3259 BP; 944 A; 695 C; 674 G; 946 T; 0 U; 0 Other;

Query Match 85.7%; Score 1793; DB 5; Length 3259;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2093; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

1 ATGAGAGAGTTCTGTACTATATGTCTACACAGCAGGAGCAGGCAAGCCATCGCAGAA 60  
80 ATGAGAGAGTTCTGTACTATATGTCTACACAGCAGGAGCAGGCAAGCCATCGCAGAA 139  
61 GAAATGTGTAGACAGCTGTGTACATGTGATTTTCTGCAAGATCTTCACTGTATTAGTAA 120  
140 GAAATGTGTAGACAGCTGTGTACATGTGATTTTCTGCAAGATCTTCACTGTATTAGTAA 199  
121 TCCGATTAAGTATGACCTTAACAAACCGAAGAGCTCTCTGTTGTTGTTGTTCTACACAG 180  
200 TCCGATTAAGTATGACCTTAACAAACCGAAGAGCTCTCTGTTGTTGTTGTTCTACACAG 259  
181 GGCACCGGAGACCCACCGACACAGCCCGCAAGTTGTTAAGGAAATACAGAACCAACA 240  
260 GGCACCGGAGACCCACCGACACAGCCCGCAAGTTGTTAAGGAAATACAGAACCAACA 319  
241 CTGCCGTTGATTTCTTGTCTCACTGCGGATAGGTTACTGGGTCTCGGTATTACAGA 300  
320 CTGCCGTTGATTTCTTGTCTCACTGCGGATAGGTTACTGGGTCTCGGTATTACAGA 379  
301 TACACCTACTTTTGGCAATGGGGGGAGATTAATTGATTAACGACTTCAAGAGCTTGAGCC 360  
380 TACACCTACTTTTGGCAATGGGGGGAGATTAATTGATTAACGACTTCAAGAGCTTGAGCC 439  
361 CGGCAATTTCTATGACACTGAGACATGAGATGACTGTAGTGTAGAACCTTGTGTGTAG 420  
440 CGGCAATTTCTATGACACTGAGACATGAGATGACTGTAGTGTAGAACCTTGTGTGTAG 499  
421 CCGTGTATGCTGTGACTCTGTGCGACAGCTTCAAGAACATTTTATGTCANAGCAGAGCA 480  
500 CCGTGTATGCTGTGACTCTGTGCGACAGCTTCAAGAACATTTTATGTCANAGCAGAGCA 559  
481 GAGGAGATTAAGTGGGCGACCTCCCGGTGGATCACTGCAATCTTGAGGACAGACCTTGG 540  
560 GAGGAGATTAAGTGGGCGACCTCCCGGTGGATCACTGCAATCTTGAGGACAGACCTTGG 619  
541 AAGTCAGAGCTGTACATATGATCTCAAGTGTGAGCTTCTGAGATTCAGATTCAGAGA 600  
620 AAGTCAGAGCTGTACATATGATCTCAAGTGTGAGCTTCTGAGATTCAGATTCAGAGA 679  
601 AAGAAAGATTTCTGAGGTTTGAAGCAAAATGCAGTGAACAGCAACCAATCCAAATGTTGA 660  
680 AAGAAAGATTTCTGAGGTTTGAAGCAAAATGCAGTGAACAGCAACCAATCCAAATGTTGA 739  
661 ATTGAAGACTTTGAGTCTCACTTACCCGTTGGTACCCCACTCTCAAGAGCTTCTG 720  
740 ATTGAAGACTTTGAGTCTCACTTACCCGTTGGTACCCCACTCTCAAGAGCTTCTG 799  
721 AATATTCCTGTGTTACCCCAAGATTTTACAGGTACATCTGAGAGAGTCTTGTGCGCAG 780  
800 AATATTCCTGTGTTACCCCAAGATTTTACAGGTACATCTGAGAGAGTCTTGTGCGCAG 859  
781 GAGGAAAGCCAGATCTGTGACTTCAAGAGATCCAGTTTTCAGATGCAATTTCAAG 840  
860 GAGGAAAGCCAGATCTGTGACTTCAAGAGATCCAGTTTTCAGATGCAATTTCAAG 919  
841 GCAATTCCTGTTACCTGATGATGATGCAATTAACCACTCTGTTGATTAATGAGCAT 900  
920 GCAATTCCTGTTACCTGATGATGATGCAATTAACCACTCTGTTGATTAATGAGCAT 979  
901 TCAATATACAGACTTTTCTATGAGCTGTGAGAGCTTCTGAGGTGATCTGCTTAACAGT 960  
980 TCAATATACAGACTTTTCTATGAGCTGTGAGAGCTTCTGAGGTGATCTGCTTAACAGT 1039  
961 GATTTGAGGTACAAAGCTTACTCAAGAGCTGAGCTTGAAGATTAAGAGAGCACTGC 1020  
1040 GATTTGAGGTACAAAGCTTACTCAAGAGCTGAGCTTGAAGATTAAGAGAGCACTGC 1099

1021 GTCTTTTGAATAAAGCAGACAAAGAAAGAGAGCTTACCTTACCCAGCATATA 1080  
1100 GTCTTTTGAATAAAGCAGACAAAGAAAGAGAGCTTACCTTACCCAGCATATA 1159  
1081 CTTGGGGATGTTCTCTCAGTTCAATTTTACCTGTGTCTTGAATCCAGCAATTTCT 1140  
1160 CTTGGGGATGTTCTCTCAGTTCAATTTTACCTGTGTCTTGAATCCAGCAATTTCT 1219  
1141 AAAAGGCAATTTTGGAGCCCTGTGACATATACAGTACAGTGTGAAAAGGCGAG 1200  
1220 AAAAGGCAATTTTGGAGCCCTGTGACATATACAGTACAGTGTGAAAAGGCGAG 1279  
1201 CTACAGAGCTGTGACATTAACAGAGGCGAGCTTATAGCCGTTGTATACAGATGCC 1260  
1280 CTACAGAGCTGTGACATTAACAGAGGCGAGCTTATAGCCGTTGTATACAGATGCC 1339  
1261 TGTGCTGTGTGTGATCTCTCTCGTTTCCCTTCTTGCCAGCCACACTCACTCTC 1320  
1340 TGTGCTGTGTGTGATCTCTCTCGTTTCCCTTCTTGCCAGCCACACTCACTCTC 1399  
1321 CTGCTCGAAATCTTCTTAACTTCAACCCAGACATATGAGTGTGCAAGCTCAAGTTA 1380  
1400 CTGCTCGAAATCTTCTTAACTTCAACCCAGACATATGAGTGTGCAAGCTCAAGTTA 1459  
1381 TTTCAACCCAGAAAGCTCCTCAATTTGTCTCAATTTGTGTAATTTCTGTCTACAGCA 1440  
1460 TTTCAACCCAGAAAGCTCCTCAATTTGTCTCAATTTGTGTAATTTCTGTCTACAGCA 1519  
1441 ACAAGGTTCTGCGAAGGAGATATGACAGCTGTGCTGTGTGTTGCTTCAAGTT 1500  
1520 ACAAGGTTCTGCGAAGGAGATATGACAGCTGTGCTGTGTGTTGCTTCAAGTT 1579  
1501 CTTGAGCGAAATATCATGATCACTCCATGAAAGCACCGGAAAGCCCTGCTCTTAAGATA 1560  
1580 CTTGAGCGAAATATCATGATCACTCCATGAAAGCACCGGAAAGCCCTGCTCTTAAGATA 1639  
1561 TCCATCTCTCTCGAACAACAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
1640 TCCATCTCTCTCGAACAACAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1699  
1621 ATATAGTGTGTGTCAGAAACCGGCAATAGCCCGTTTATTTGGTTCTTCAACAT---AG 1676  
1700 ATATAGTGTGTGTCAGAAACCGGCAATAGCCCGTTTATTTGGTTCTTCAACATAGAGAG 1759  
1677 AAATCTCAAGAACCAACCCAGATGGAATTTTGAAGAACATGATGTTTGTGCTG 1736  
1760 AAATCTCAAGAACCAACCCAGATGGAATTTTGAAGAACATGATGTTTGTGCTG 1819  
1737 AGGCAATAGGAGATTAATCTATTCAGAAAGAGCTCAGACATTTCTTAAGCATGGG 1796  
1820 AGGCAATAGGAGATTAATCTATTCAGAAAGAGCTCAGACATTTCTTAAGCATGGG 1879  
1797 ATCTTAATCTCATTAAGGTTTCTTCTCAAGATGCTCTGTGTGGAGAGAGAGCC 1856  
1880 ATCTTAATCTCATTAAGGTTTCTTCTCAAGATGCTCTGTGTGGAGAGAGAGCC 1939  
1857 CCAAGAAATATGATACAGAACATCCAGCTTCACTGAGCAGAGGTGCGAGAAATCTCTC 1916  
1940 CCAAGAAATATGATACAGAACATCCAGCTTCACTGAGCAGAGGTGCGAGAAATCTCTC 1999  
1917 CTCCAGAGAGAGGCGCATATTTATGTGTGTGAGATGCAAGAAATATGAGCAAGATGA 1976  
2000 CTCCAGAGAGAGGCGCATATTTATGTGTGTGAGATGCAAGAAATATGAGCAAGATGA 2059  
1977 CATGATGCCCTTGTGCAATAAAGCAAGAGGTTGAGTTGAAAATCTAAGCAATG 2036  
2060 CATGATGCCCTTGTGCAATAAAGCAAGAGGTTGAGTTGAAAATCTAAGCAATG 2119  
2037 AAAACCTGTGCACTTTTAAAGAAAGAAAGCTTACCTTCAAGATTTTGTGCTATTA 2093  
2120 AAAACCTGTGCACTTTTAAAGAAAGAAAGCTTACCTTCAAGATTTTGTGCTATTA 2176



RESULT 4  
ADMA3206  
ID ADMA3206 standard; cDNA; 3259 BP.  
XX  
AC ADMA3206;  
XX  
DT 03-JUN-2004 (first entry)  
XX  
DE Human full length cDNA encoding methionine synthase reductase.  
XX  
KW Human; ss; gene; Methionine synthase reductase polypeptide; HAMTR; cancer; cardiovascular disease; neural tube defect; hyperhomocysteinemia; chromosome 5p15.2-p15.3; SNP;  
KM single nucleotide polymorphism.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 80..2176  
FT FT /\*tag= a  
FT FT /product= "hamTRR"  
FT FT replace(145, A)  
FT FT /\*tag= b  
FT FT /standard\_name= "Single\_nucleotide\_polymorphism"  
FT FT replace(189, A)  
FT FT /\*tag= c  
FT FT /standard\_name= "Single\_nucleotide\_polymorphism"  
XX  
XX US2003082676-A1.  
XX  
XX 01-MAY-2003.  
XX  
XX 10-AUG-1999; 99US-00371347.  
XX  
XX 16-JAN-1998; 98US-0071622P.  
XX PR 15-JAN-1999; 99US-00232028.  
XX  
XX (GRAY/) GRAVEL R A.  
XX PA (ROZE/) ROZEN R.  
XX PA (LECL/) LECLERC D.  
XX PA (WILS/) WILSON A.  
XX PA (ROSE/) ROSENBLATT D.  
XX  
XX PI Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;  
XX DR WPI; 2003-576610/54.  
XX DR P-PSDB; ADMA3207.  
XX  
XX PT New substantially pure nucleic acid encoding a mammalian methionine  
XX PT synthase reductase polypeptide, useful for diagnosing, preventing or  
XX PT treating conditions associated with altered methionine synthase activity,  
XX PT e.g. cancer.  
XX  
XX PS Example 2; SEQ ID NO 24; 26pp; English.  
XX  
XX The invention relates to a substantially pure nucleic acid that encodes a  
XX CC mammalian methionine synthase reductase polypeptide, HAMTR, or that  
XX CC hybridizes at high stringency to a nucleic acid appearing as ADMA3208 or  
XX CC ADMA3209. Also included are a non-human animal where one or both genetic  
XX CC alleles encoding the methionine synthase reductase polypeptide are  
XX CC mutated, an antibody that specifically binds the above methionine  
XX CC synthase reductase polypeptide, a method of detecting the presence of the  
XX CC above polypeptide, a method for detecting sequence variants for  
XX CC methionine synthase reductase in a mammal, methods of treating or  
XX CC preventing cancer (or cardiovascular disease or neural tube defects) in a  
XX CC subject, methods of screening for a compound that modulates methionine  
XX CC synthase reductase biological activity and a method for detecting an  
XX CC increased risk of developing a neural tube defect in a mammalian embryo  
XX CC or foetus. The nucleic acid is useful in diagnosing, preventing or  
XX CC treating conditions associated with altered methionine synthase activity,  
XX CC such as cancer, cardiovascular disease or neural tube defects, or in  
XX CC screening for a compound that modulates methionine synthase reductase  
XX CC biological activity. Naturally occurring variants of the polypeptide are

CC also associated with hyperhomocysteinemia. The gene for HAMTR is  
CC located on chromosome 5p15.2-p15.3. The present sequence is full length  
CC sequence of the wild-type human hamTRR cDNA.  
XX  
SQ Sequence 3259 BP; 944 A; 695 C; 674 G; 946 T; 0 U; 0 Other;  
Query Match 85.7%; Score 1793; DB 11; Length 3259;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2093; Conservative 0; Mismatches 0; Indels 4; Gaps 1;  
QY 1 ATGAGAGGTTTCTGTATCAATATGCTACACAGCAGGACAGGCAAGCCATCGAGAA 60  
Db 80 ATGAGAGGTTTCTGTATCAATATGCTACACAGCAGGACAGGCAAGCCATCGAGAA 139  
QY 61 GAAATGTGAGCAAGCTGTGTACATGATTTTCTGCAGATCTTCACTGATTAAGTAA 120  
Db 140 GAAATGTGAGCAAGCTGTGTACATGATTTTCTGCAGATCTTCACTGATTAAGTAA 199  
QY 121 TCCGATTAAGTAACTTAATAAACCAGAACAGCTCCCTGTTGTTGTTTCTACACG 180  
Db 200 TCCGATTAAGTAACTTAATAAACCAGAACAGCTCCCTGTTGTTGTTTCTACACG 259  
QY 181 GGCACCGAGACCCACCCGACACAGCCCGCAGATTGTTAAGAAATACAGAACAAACA 240  
Db 260 GGCACCGAGACCCACCCGACACAGCCCGCAGATTGTTAAGAAATACAGAACAAACA 319  
QY 241 CTGCGGTTGATTTCTTTGCTACCTGCGGATGAGTTACTGGGTTCCGTTATTCAGAA 300  
Db 320 CTGCGGTTGATTTCTTTGCTACCTGCGGATGAGTTACTGGGTTCCGTTATTCAGAA 379  
QY 301 TACACCTACTTTTGAATGGGGGAGATTAATGATTAACGACTTCAAGAGCTTGAGGC 360  
Db 380 TACACCTACTTTTGAATGGGGGAGATTAATGATTAACGACTTCAAGAGCTTGAGGC 439  
QY 361 CGGCATTTCTATGACACTGACATGACATGACTGTGTAGTTTAAACTTGTGTGAG 420  
Db 440 CGGCATTTCTATGACACTGACATGACATGACTGTGTAGTTTAAACTTGTGTGAG 499  
QY 421 CCGTGATTTCTGAGACTGTGGCCAGCTTCAGAACATTTTATGTCAGAGAGACAA 480  
Db 500 CCGTGATTTCTGAGACTGTGGCCAGCTTCAGAACATTTTATGTCAGAGAGACAA 559  
QY 481 GAGGATTAAGTGGCCGACCTCCCGTGGCATCCGTCGATCCCTTGAGGACAGACCTTGG 540  
Db 560 GAGGATTAAGTGGCCGACCTCCCGTGGCATCCGTCGATCCCTTGAGGACAGACCTTGG 619  
QY 541 AAGTCAGAGCTGTACACATTTGATCTCAAGTCGAGCTTCTGAGATTCTGATTCAGGA 600  
Db 620 AAGTCAGAGCTGTACACATTTGATCTCAAGTCGAGCTTCTGAGATTCTGATTCAGGA 679  
QY 601 AGAAGGATTTGAGGTTTGAAGCAAAATGCAGTGAACAGCAACCAATCCAAATGTTGTA 660  
Db 680 AGAAGGATTTGAGGTTTGAAGCAAAATGCAGTGAACAGCAACCAATCCAAATGTTGTA 739  
QY 661 ATTGAAGCTTGAAGCTTCACCTTACCCGTTGGTATCCGCCATCTCACAGGCTCTCG 720  
Db 740 ATTGAAGCTTGAAGCTTCACCTTACCCGTTGGTATCCGCCATCTCACAGGCTCTCG 799  
QY 721 AATATTCCTGTTTATCCCCAGATATTTTACAGATCATCTGACAGAGTCTTGTGCCAG 780  
Db 800 AATATTCCTGTTTATCCCCAGATATTTTACAGATCATCTGACAGAGTCTTGTGCCAG 859  
QY 781 GAGAAAGCCAGATATCTGTGACTTCAGAGATTCAGATTTTCAAGTCCAAATTCAGAG 840  
Db 860 GAGAAAGCCAGATATCTGTGACTTCAGAGATTCAGATTTTCAAGTCCAAATTCAGAG 919  
QY 841 GCAGTCAACTTATCTGAGATATGCAATTAACCAAGCTGCTGTAGATTTGAGCATTT 900  
Db 920 GCAGTCAACTTATCTGAGATATGCAATTAACCAAGCTGCTGTAGATTTGAGCATTT 979  
QY 901 TCAATATCAGACTTTTCTATCAGCTTGAGATGCTTGAAGCTGTATGCTTAAACAGT 960  
Db 980 TCAATATCAGACTTTTCTATCAGCTTGAGATGCTTGAAGCTGTATGCTTAAACAGT 1039

QY 961 GATTCGAGGTACAAAGCCTACTCCAAAGACTGACGCTTGAAGATTAAGAGAGCACTGC 1020  
 DB 1040 GATTCGAGGTACAAAGCCTACTCCAAAGACTGACGCTTGAAGATTAAGAGAGCACTGC 1099  
 QY 1021 GTCTCTTTGAAAAATTAAGGACAGACAAAGAAAGAAAGAGCTTACTTACCCGACATATA 1080  
 DB 1100 GTCTCTTTGAAAAATTAAGGACAGACAAAGAAAGAAAGAGCTTACTTACCCGACATATA 1159  
 QY 1081 CCTGCGGGAGTGTCTCTCCAGTTCAATTTTACCTGCTGTCTGAATCCGAGCAATTCCT 1140  
 DB 1160 CCTGCGGGAGTGTCTCTCCAGTTCAATTTTACCTGCTGTCTGAATCCGAGCAATTCCT 1219  
 QY 1141 AAAAAGGCAATTTTGGAGCCCTTGTGAGCTATACAGTGAAGCTGTGAAAAAGGACAG 1200  
 DB 1220 AAAAAGGCAATTTTGGAGCCCTTGTGAGCTATACAGTGAAGCTGTGAAAAAGGACAG 1279  
 QY 1201 CTACAGAGCTGTGACGTAAACAAAGGGGAGCCGATTAATAGCCGCTTGTGACGAGATGCC 1260  
 DB 1280 CTACAGAGCTGTGACGTAAACAAAGGGGAGCCGATTAATAGCCGCTTGTGACGAGATGCC 1339  
 QY 1261 TGTGCTGCTGTGTGGATCTCTCTGCTTCTTCTTCCGACGACCACTCACTCTC 1320  
 DB 1340 TGTGCTGCTGTGTGGATCTCTCTGCTTCTTCTTCCGACGACCACTCACTCTCTC 1399  
 QY 1321 CTGCTGGAACATCTTCTTAACTTCAACCCAGACATATTGCTGACAGCTCAAGTTTA 1380  
 DB 1400 CTGCTGGAACATCTTCTTAACTTCAACCCAGACATATTGCTGACAGCTCAAGTTTA 1459  
 QY 1381 TTTCAACCAAGAAAGCTCCATTTTGTCTTCAACATTTGGAATTTCTGTCTACTGACAA 1440  
 DB 1460 TTTCAACCAAGAAAGCTCCATTTTGTCTTCAACATTTGGAATTTCTGTCTACTGACAA 1519  
 QY 1441 ACAAGGTTCTGCGGAAGGAGATATGACAGCTGCTGCTGCTTGTGTGCTTCACTT 1500  
 DB 1520 ACAAGGTTCTGCGGAAGGAGATATGACAGCTGCTGCTGCTTGTGTGCTTCACTT 1579  
 QY 1501 CTTTCAACCAAGCAATATGATGATCCCATGAAAGACAGGGGAAAGCCCTGGCTCTTAAGATA 1560  
 DB 1580 CTTTCAACCAAGCAATATGATGATCCCATGAAAGACAGGGGAAAGCCCTGGCTCTTAAGATA 1639  
 QY 1561 TCCATCTCTCTCTGGAACAACAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
 DB 1640 TCCATCTCTCTCTGGAACAACAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1699  
 QY 1621 ATATATGTGGGTCCAGGAACCGGACATAGCCCTTATTTGGGTTCTTACAACT---AG 1676  
 DB 1700 ATATATGTGGGTCCAGGAACCGGACATAGCCCTTATTTGGGTTCTTACAACTAGAGAG 1759  
 QY 1677 AAATCTCAAGAAACAACAACCCAGATGGAATTTTGGAGCAATGTGTGTTTTTGGCTGC 1736  
 DB 1760 AAATCTCAAGAAACAACAACCCAGATGGAATTTTGGAGCAATGTGTGTTTTTGGCTGC 1819  
 QY 1737 AGGATATAGAGATAGGATTAATCTAATCAGAAAAGAGCTCAGACATTTCTTAAGCATGGG 1796  
 DB 1820 AGGATATAGAGATAGGATTAATCTAATCAGAAAAGAGCTCAGACATTTCTTAAGCATGGG 1879  
 QY 1797 ATCTTAATCTAATTAAGGTTCTTCTCAAGAGATGCTCTGTGTGGGAGAGAGAGGCC 1856  
 DB 1880 ATCTTAATCTAATTAAGGTTCTTCTCAAGAGATGCTCTGTGTGGGAGAGAGAGGCC 1939  
 QY 1857 CAGGAAAGATATAGAGCAACAATCCAGCTTCAATGAGGAGAGAGAGAGAGAGAGAGAGAG 1916  
 DB 1940 CAGGAAAGATATAGAGCAACAATCCAGCTTCAATGAGGAGAGAGAGAGAGAGAGAGAGAG 1999  
 QY 1917 CTCACAGAGACGGCCATATTTATGTGTGTGAGATGCAAGAAATATGCGCAAGATGTA 1976  
 DB 2000 CTCACAGAGAGACGGCCATATTTATGTGTGTGAGATGCAAGAAATATGCGCAAGATGTA 2059  
 QY 1977 CATGATGCTTGTGCAATATATAGCAAAAGAGTTGAGTTGAAAACTAGAGCAATG 2036  
 DB 2060 CATGATGCTTGTGCAATATATAGCAAAAGAGTTGAGTTGAAAACTAGAGCAATG 2119

QY 2037 AAAACCTGGCCACTTTAAAGAGAAAAACGCTTACTTACAGATATTTGTCTATTA 2093  
 DB 2120 AAAACCTGGCCACTTTAAAGAGAAAAACGCTTACTTACAGATATTTGTCTATTA 2176  
 RESULT 5  
 ID AD43208  
 AD43208 standard; cDNA; 2094 BP.  
 XX  
 AC AD43208;  
 XX  
 DT 03-JUN-2004 (first entry)  
 XX  
 DE Human wild-type methionine synthase reductase CDS.  
 XX  
 KW Human; ss; Methionine synthase reductase polypeptide; hsmTRR; cancer;  
 KW cardiovascular disease; neural tube defect; hyperhomocysteinemia;  
 KW chromosome 5p15.2-p15.3; SNP; single nucleotide polymorphism.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..2094  
 FT /\*tag= a  
 FT /product= "hsmTRR"  
 FT /partial  
 FT /note= "No stop codon shown"  
 FT replace(66,A)  
 FT /\*tag= b  
 FT /standard\_name= "single\_nucleotide\_polymorphism"  
 FT replace(110,A)  
 FT /\*tag= c  
 FT /standard\_name= "single\_nucleotide\_polymorphism"  
 XX  
 EN US2003082676-A1.  
 XX  
 PD 01-MAY-2003.  
 XX  
 PF 10-AUG-1999; 99US-00371347.  
 XX  
 PR 16-JAN-1998; 98US-0071622P.  
 PR 15-JAN-1999; 99US-00232028.  
 XX  
 PA (GRAY/) GRAVEL R. A.  
 PA (ROZE/) ROZEN R. R.  
 PA (LECL/) LECLERC D.  
 PA (WILS/) WILSON A.  
 PA (ROSE/) ROSENBLATT D.  
 XX  
 PI Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;  
 XX  
 DR WPI; 2003-576610/54.  
 DR P-PSDB; AD43207.  
 XX  
 PT New substantially pure nucleic acid encoding a mammalian methionine  
 PT synthase reductase polypeptide, useful for diagnosis, preventing or  
 PT treating conditions associated with altered methionine synthase activity,  
 PT e.g. cancer.  
 XX  
 PS Claim 3; SEQ ID NO 1; 26pp; English.  
 XX  
 CC The invention relates to a substantially pure nucleic acid that encodes a  
 CC mammalian methionine synthase reductase polypeptide, hsmTRR, or that  
 CC hybridizes at high stringency to a nucleic acid appearing as AD43208 or  
 CC AD43209. Also included are a non-human animal where one or both genetic  
 CC alleles encoding the methionine synthase reductase polypeptide are  
 CC mutated, an antibody that specifically binds the above methionine  
 CC synthase reductase polypeptide, a method of detecting the presence of the  
 CC above polypeptide, a method for detecting sequence variants for  
 CC methionine synthase reductase in a mammal, methods of treating or  
 CC preventing cancer (or cardiovascular disease or neural tube defects) in a  
 CC subject, methods of screening for a compound that modulates methionine  
 CC synthase reductase biological activity and a method for detecting an



DB 1921 CTCGAGAGAGAGCCATATTTATGTGTGTGATGCAAGAAATATGCGCAAGATGTA 1980  
QY 1977 CATTGATCCCTTGTGCAATTAATTAAGCAAGAGTTGAGTTGAAAACTAGAGCAATG 2036  
DB 1981 CATGATGCCCCCTTGTGCAATTAATTAAGCAAGAGTTGAGTTGAAAACTAGAGCAATG 2040  
QY 2037 AAAACCTGCGCACTTTAAAAAGAAAGAAACGTTACCTTCAGGATATTTGGTCA 2090  
DB 2041 AAAACCTGCGCACTTTAAAAAGAAAGAAAGAAACGTTACCTTCAGGATATTTGGTCA 2094

RESULT 6  
ADM43212  
ID ADM43212 standard; cDNA; 2094 BP.  
XX ADM43212;  
AC  
XX  
XX 03-JUN-2004 (first entry)  
DE Human methionine synthase reductase CDS G110A variant.  
XX  
XX Human; ss; Methionine synthase reductase polypeptide; HmMTR; cancer;  
KM cardiovascular disease; neural tube defect; hyperhomocysteinemia;  
XX chromosome 5p15.2-p15.3; SNP; single nucleotide polymorphism.  
OS Homo sapiens.

XX  
XX  
XX Key Location/Qualifiers  
FT CDS 1..2094  
FT /tag= a  
FT /product= "HmMTR"  
FT /partial  
FT /note= "No stop codon shown"  
FT variation replace (66,A)  
FT /standard\_name= "Single\_nucleotide\_polymorphism"  
FT /tag= C  
FT /standard\_name= "Single\_nucleotide\_polymorphism"

XX  
XX US2003082676-A1.  
XX  
XX 01-MAY-2003.  
XX  
XX 10-AUG-1999; 99US-00371347.  
XX  
XX 16-JAN-1998; 98US-0071622P.  
XX 15-JAN-1999; 99US-00232028.  
XX  
XX (GRAY/) GRAVEL R A.  
XX (ROZE/) ROZEN R.  
XX (LECL/) LECLEERC D.  
XX (WILS/) WILSON A.  
XX (ROSE/) ROSENBLATT D.  
XX  
XX Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;  
XX  
XX MPI; 2003-576610/54.  
XX P-PSDB; ADM43213.  
XX  
XX  
XX New substantially pure nucleic acid encoding a mammalian methionine  
PT synthase reductase polypeptide, useful for diagnosing, preventing or  
PT treating conditions associated with altered methionine synthase activity,  
PT e.g. cancer.  
XX  
XX  
XX Disclosure; SEQ ID NO 43; 26pp; English.

XX  
XX The invention relates to a substantially pure nucleic acid that encodes a  
CC mammalian methionine synthase reductase polypeptide, HmMTR, or that  
CC hybridizes at high stringency to a nucleic acid appearing as ADM43208 or  
CC ADM43209. Also included are a non-human animal where one or both genetic  
CC alleles encoding the methionine synthase reductase polypeptide are  
CC mutated, an antibody that specifically binds the above methionine

CC synthase reductase polypeptide, a method of detecting the presence of the  
CC above polypeptide, a method for detecting sequence variants for  
CC methionine synthase reductase in a mammal, methods of treating or  
CC preventing cancer (or cardiovascular disease or neural tube defects) in a  
CC subject, methods of screening for a compound that modulates methionine  
CC synthase reductase biological activity and a method for detecting an  
CC increased risk of developing a neural tube defect in a mammalian embryo  
CC or foetus. The nucleic acid is useful in diagnosing, preventing or  
CC treating conditions associated with altered methionine synthase activity,  
CC such as cancer, cardiovascular disease or neural tube defects, or in  
CC screening for a compound that modulates methionine synthase reductase  
CC biological activity. Naturally occurring variants of the polypeptide are  
CC also associated with hyperhomocysteinemia. The gene for HmMTR is  
CC located on chromosome 5p15.2-p15.3. The present sequence is the coding  
CC sequence of a variant human HmMTR cDNA.  
XX  
XX Sequence 2094 BP; 592 A; 489 C; 480 G; 533 T; 0 U; 0 Other;  
SQ

Query Match 83.1%; Score 1739; DB 11; Length 2094;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2089; Conservative 0; Mismatches 1; Indels 4; Gaps 1;

QY 1 ATGAGAGGTTTCTGTACTATATGCTACACAGCGGACAGGCAAGCCATCGCAGAA 60  
DB 1 ATGAGAGGTTTCTGTACTATATGCTACACAGCGGACAGGCAAGCCATCGCAGAA 60  
QY 61 GAAATGTGTGACCAAGCTGTGTACATGATTTTCGAGATCTTCAGTATTAAGAA 120  
DB 61 GAAATGTGTGACCAAGCTGTGTACATGATTTTCGAGATCTTCAGTATTAAGAA 120  
QY 121 TCGATATGATATGACCTTAACCAAGCAAGCTCTCTGTTGTTGTTCTTACACG 180  
DB 121 TCGATATGATATGACCTTAACCAAGCAAGCTCTCTGTTGTTGTTCTTACACG 180  
QY 181 GGCACCGAGAGACCCACCGACACAGCCCGCAAGTTGTTAAGAAATACAGAAACAA 240  
DB 181 GGCACCGAGAGACCCACCGACACAGCCCGCAAGTTGTTAAGAAATACAGAAACAA 240  
QY 241 CTGCCGTTGATTTCTTGTCTCACTCGGTATGAGTTTCTGGTCTCGGTATTCAGAA 300  
DB 241 CTGCCGTTGATTTCTTGTCTCACTCGGTATGAGTTTCTGGTCTCGGTATTCAGAA 300  
QY 301 TACACCTACTTTTGTGCAATGCGGGGAGATATGATTAACGACTTCAAGAGCTTGAGCC 360  
DB 301 TACACCTACTTTTGTGCAATGCGGGGAGATATGATTAACGACTTCAAGAGCTTGAGCC 360  
QY 361 CGGCAATTTCTATGACACTGACATGACATGACATGACATGACATGACATGACATGAC 420  
DB 361 CGGCAATTTCTATGACACTGACATGACATGACATGACATGACATGACATGACATGAC 420  
QY 421 CCGTGAATTCGTGATCTGTGCGCAGCCCTCAGAAAGCAATTTTAGTCAAGACAGACAA 480  
DB 421 CCGTGAATTCGTGATCTGTGCGCAGCCCTCAGAAAGCAATTTTAGTCAAGACAGACAA 480  
QY 481 GAGAGATATAGTGGGCACTCCCGGTGATCATCTGATCTTGTGAGACAGACTTTG 540  
DB 481 GAGAGATATAGTGGGCACTCCCGGTGATCATCTGATCTTGTGAGACAGACTTTG 540  
QY 541 AAGTCAGAGCTGTACATCAATGATCTCAAGTCACTTCAAGATTCAGATTCAGAG 600  
DB 541 AAGTCAGAGCTGTACATCAATGATCTCAAGTCACTTCAAGATTCAGATTCAGAG 600  
QY 601 AGAAAGATTCGAGTTTGAAGCAAAATGCAATGCAAGCAACCAATTCATGTTGTA 660  
DB 601 AGAAAGATTCGAGTTTGAAGCAAAATGCAATGCAAGCAACCAATTCATGTTGTA 660  
QY 661 ATTAAGATTCGAGTCTCACTTACCGTGGTACCCCACTTCACAGAGCTCTG 720  
DB 661 ATTAAGATTCGAGTCTCACTTACCGTGGTACCCCACTTCACAGAGCTCTG 720  
QY 721 AATATTCCTGTTTACCCCAAGAAATATTAAGATTCATCTGACAGAGTCTTGGCCAG 780  
DB 721 AATATTCCTGTTTACCCCAAGAAATATTAAGATTCATCTGACAGAGTCTTGGCCAG 780

QY 781 GAGAAAGCCAGATCTGTGACTTCAGCAGATCCAGTTTTCAGAGTCCCAATTTCCAAAG 840  
Db 781 GAGAAAGCCAGATCTGTGACTTCAGCAGATCCAGTTTTCAGAGTCCCAATTTCCAAAG 840  
QY 841 GCAGTTCAATTTACTAGATGATGCAATTAACCACTCTGCTGTGATTAATTTGACATTT 900  
Db 841 GCAGTTCAATTTACTAGATGATGCAATTAACCACTCTGCTGTGATTAATTTGACATTT 900  
QY 901 TCAAAATCAGACTTTTCTCTATTCAGCCTGAGATGCTTCAGCCTGATCTCCCTTAACAT 960  
Db 901 TCAAAATCAGACTTTTCTCTATTCAGCCTGAGATGCTTCAGCCTGATCTCCCTTAACAT 960  
QY 961 GATTCTGAGGTACCAAGCCTTACTCCAAAGACTGCAAGTTGAAGTAAAGAGACACTGC 1020  
Db 961 GATTCTGAGGTACCAAGCCTTACTCCAAAGACTGCAAGTTGAAGTAAAGAGACACTGC 1020  
QY 1021 GTGCTTTTGAATTAAGAGAGACACAAAGAAAGAAAGACTCTCCCTTACCCGACATATA 1080  
Db 1021 GTGCTTTTGAATTAAGAGAGACACAAAGAAAGAAAGACTCTCCCTTACCCGACATATA 1080  
QY 1081 CCTGCGGAGATGTTCTCTCAGATTCATTTTACCTGTGTCTGTAATCCGAGCAATTCCT 1140  
Db 1081 CCTGCGGAGATGTTCTCTCAGATTCATTTTACCTGTGTCTGTAATCCGAGCAATTCCT 1140  
QY 1141 AAAAAGGCATTTTGGAGCCCTTGTGACTATACAGTGAACAGTGTGAAAAGCCGACG 1200  
Db 1141 AAAAAGGCATTTTGGAGCCCTTGTGACTATACAGTGAACAGTGTGAAAAGCCGACG 1200  
QY 1201 CTACAGAGCTGTGACGTAAACAAAGAGGAGCCGATTAATGACCGCTTTGACAGATGCC 1260  
Db 1201 CTACAGAGCTGTGACGTAAACAAAGAGGAGCCGATTAATGACCGCTTTGACAGATGCC 1260  
QY 1261 TGTGCTGTGTTGTGATCTCTCTCTGCTTCCCTCTTCCAGCAGCCAGCTCAGTCTC 1320  
Db 1261 TGTGCTGTGTTGTGATCTCTCTCTGCTTCCCTCTTCCAGCAGCCAGCTCAGTCTC 1320  
QY 1321 CTGCTGGAACATCTTCTTAACTTCAACCCAGAACATATTCGTGTGACAGCTCAAGTTTA 1380  
Db 1321 CTGCTGGAACATCTTCTTAACTTCAACCCAGAACATATTCGTGTGACAGCTCAAGTTTA 1380  
QY 1381 TTTTCAACCCAGAAAGCCATTTTGTCTGACAACTGTGAAATTTCTGTCTACGCGACA 1440  
Db 1381 TTTTCAACCCAGAAAGCCATTTTGTCTGACAACTGTGAAATTTCTGTCTACGCGACA 1440  
QY 1441 ACAGAGTTCGTGCGAGAGGAGATATGACAGCTGTGCTGTGTGTTGCTTCAAGTT 1500  
Db 1441 ACAGAGTTCGTGCGAGAGGAGATATGACAGCTGTGCTGTGTGTTGCTTCAAGTT 1500  
QY 1501 CTTGACGCCAAACATATCATGATCCATGAAAGACAGCGGAAAGCCCTGCTCTAAGATA 1560  
Db 1501 CTTGACGCCAAACATATCATGATCCATGAAAGACAGCGGAAAGCCCTGCTCTAAGATA 1560  
QY 1561 TCCATCTCTCTCTGGAACAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
Db 1561 TCCATCTCTCTCTGGAACAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
QY 1621 ATATATGTGGGTCCAGGAAACCGGATATGCCCCGTTTATTTGGGTTCTTCAACATAGAG 1680  
Db 1621 ATATATGTGGGTCCAGGAAACCGGATATGCCCCGTTTATTTGGGTTCTTCAACATAGAG 1680  
QY 1677 AAACTCCAGAAACACCCAGATGGAATTTTGGAGCAATGTGTGTTTGTGCTGC 1736  
Db 1677 AAACTCCAGAAACACCCAGATGGAATTTTGGAGCAATGTGTGTTTGTGCTGC 1736  
QY 1737 AGGATTAAGATAGGAGTTATCTATTCAGAAAAGAGCTCAGATTTCTTAAAGCATGG 1796  
Db 1741 AGGATTAAGATAGGAGTTATCTATTCAGAAAAGAGCTCAGATTTCTTAAAGCATGG 1800  
QY 1797 ATCTTAATCATCTAAAGGTTTCTTCTCAAGAGATCTCTGTGTGGAGAGAGAGCC 1856  
Db 1801 ATCTTAATCATCTAAAGGTTTCTTCTCAAGAGATCTCTGTGTGGAGAGAGAGCC 1860

QY 1857 CCAGCAAGTATGATCAAGACAAATCATCAGCTTCATGGCCAGGAGGTGGCAGAAATCTC 1916  
Db 1861 CCAGCAAGTATGATCAAGACAAATCATCAGCTTCATGGCCAGGAGGTGGCAGAAATCTC 1920  
QY 1917 CTCACAGAAAGCCGCAATTTATTTGTGTGTGAGATGCAAAAGATATGCGCAAGATGTA 1976  
Db 1921 CTCACAGAAAGCCGCAATTTATTTGTGTGTGAGATGCAAAAGATATGCGCAAGATGTA 1980  
QY 1977 CATGATGCCCTTGTGCAAAATTAATTAAGCAAAAGAGTTGAGTTGAAAACTAGAAGCATG 2036  
Db 1981 CATGATGCCCTTGTGCAAAATTAATTAAGCAAAAGAGTTGAGTTGAAAACTAGAAGCATG 2040  
QY 2037 AAAACCTGGCCACTTTAAAGAAAGAAACGCTTACCTTCAGATATTTTGCTCA 2090  
Db 2041 AAAACCTGGCCACTTTAAAGAAAGAAACGCTTACCTTCAGATATTTTGCTCA 2094

RESULT 7  
ADM43209  
ID ADM43209 standard; cDNA; 2094 BP.  
XX  
AC ADM43209;  
XX  
DT 03-JUN-2004 (first entry)  
XX  
DE Human methionine synthase reductase CDS G66A variant.  
XX  
KW Human; ss; Methionine synthase reductase polypeptide; HsmTRR; cancer;  
KW cardiovascular disease; neural tube defect; hyperhomocysteinemia;  
KW chromosome 5p15.2-p15.3; SNP; single nucleotide polymorphism.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..2094  
FT /tag= a  
FT /product= "hsmTRR"  
FT /partial  
FT /note= "No stop codon shown"  
FT variation  
FT /tag= b  
FT /tag= d  
FT /standard name= "single\_nucleotide polymorphism"  
FT /replace (110,A)  
FT /tag= c  
FT /standard\_name= "single\_nucleotide polymorphism"  
XX  
PN US2003082676-A1.  
XX  
PD 01-MAY-2003.  
XX  
PE 10-AUG-1999; 99US-00371347.  
XX  
PF 16-JAN-1998; 98US-0071622P.  
PR 15-JAN-1999; 99US-00232026.  
XX  
PA (GRAV/) GRAVEL R A.  
PA (ROZE/) ROZEN R.  
PA (LECL/) LECLERC D.  
PA (WILS/) WILSON A.  
PA (ROSE/) ROSENBLATT D.  
XX  
PI Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;  
XX  
DR WPI; 2003-576610/54.  
XX  
DR P-PDB; ADM43211.  
XX  
PT New substantially pure nucleic acid encoding a mammalian methionine  
PT synthase reductase polypeptide, useful for diagnosing, preventing or  
PT treating conditions associated with altered methionine synthase activity,  
XX e.g. cancer.  
XX  
PS Claim 3; SEQ ID NO 41; 26bp; English.  
XX

The invention relates to a substantially pure nucleic acid that encodes a mammalian methionine synthase reductase polypeptide, hSMTR, or that hybridizes at high stringency to a nucleic acid appearing as AD43208 or AD43209. Also included are a non-human animal where one or both genetic alleles encoding the methionine synthase reductase polypeptide are mutated, an antibody that specifically binds the above methionine synthase reductase polypeptide, a method of detecting the presence of the above polypeptide, a method for detecting sequence variants for methionine synthase reductase in a mammal, methods of treating or preventing cancer (or cardiovascular disease or neural tube defects) in a subject, methods of screening for a compound that modulates methionine synthase reductase biological activity and a method for detecting an increased risk of developing a neural tube defect in a mammalian embryo or foetus. The nucleic acid is useful in diagnosing, preventing or treating conditions associated with altered methionine synthase activity, such as cancer, cardiovascular disease or neural tube defects, or in screening for a compound that modulates methionine synthase reductase biological activity. Naturally occurring variants of the polypeptide are also associated with hyperhomocysteinaemia. The gene for hSMTR is located on chromosome 5p15.2-p15.3. The present sequence is the coding sequence of a variant human hSMTR cDNA.

Sequence 2094 BP; 592 A; 489 C; 480 G; 533 T; 0 U; 0 Other;

Query Match 83.1%; Score 1739; DB 11; Length 2094;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2089; Conservative 0; Mismatches 1; Indels 4; Gaps 1;

```
QY 1 ATGAGAGAGTTCTGTATCTATATCTACAGAGGAGCAGCAAGCCATCCGAGAA 60
DB 1 ATGAGAGAGTTCTGTATCTATATCTACAGAGGAGCAGCAAGCCATCCGAGAA 60
QY 61 GAAATGTGTGAGCAGCTGTGTACATGATTTTCTGAGATCTTCACTGTATTAGTAA 120
DB 61 GAAATGTGTGAGCAGCTGTGTACATGATTTTCTGAGATCTTCACTGTATTAGTAA 120
QY 121 TCCGATAGTATGACCTAAACCCGAAACAGCTCTGTGTGTGTGTCTACACAG 180
DB 121 TCCGATAGTATGACCTAAACCCGAAACAGCTCTGTGTGTGTGTCTACACAG 180
QY 181 GGCACCGAGAGCCACCCGACAGCCGCAAGTTGTTAGAAATACAGAACCAACA 240
DB 181 GGCACCGAGAGCCACCCGACAGCCGCAAGTTGTTAGAAATACAGAACCAACA 240
QY 241 CTGCGCGTGTGATTTCTTGTCTCACTGCGGTATGCGTTACTGGTCTCGGTATTCAGAA 300
DB 241 CTGCGCGTGTGATTTCTTGTCTCACTGCGGTATGCGTTACTGGTCTCGGTATTCAGAA 300
QY 301 TACACCTACTTTGCAATGGGGGGAAGATATTGATTAACGACTTCAAGAGCTTGAGCC 360
DB 301 TACACCTACTTTGCAATGGGGGGAAGATATTGATTAACGACTTCAAGAGCTTGAGCC 360
QY 361 CGGCAATTTCTATGACATGAGATGACATGACTGTGATTTAGAACTTGTGTGAG 420
DB 361 CGGCAATTTCTATGACATGAGATGACATGACTGTGATTTAGAACTTGTGTGAG 420
QY 421 CCGTGTGATTTGACATCTGCGCAGCCCTCAGAAAGCATTTTAAAGTCAAGAGAGCAA 480
DB 421 CCGTGTGATTTGACATCTGCGCAGCCCTCAGAAAGCATTTTAAAGTCAAGAGAGCAA 480
QY 481 GAGGAGATTAAGTGGGCACTCCGGTGGATCACTGATCTTGAAGAGACCTTGTG 540
DB 481 GAGGAGATTAAGTGGGCACTCCGGTGGATCACTGATCTTGAAGAGACCTTGTG 540
QY 541 AAGTCAGAGCTCTACATGATTTCAAGTCTGAGATTTCAAGATTTCAAGATTTCAAG 600
DB 541 AAGTCAGAGCTCTACATGATTTCAAGTCTGAGATTTCAAGATTTCAAGATTTCAAG 600
QY 601 AAGAAAGATTTCTGAGGTTTGAAGCAAAATGCAATGAGCAACCAATTCATGTTGTA 660
DB 601 AAGAAAGATTTCTGAGGTTTGAAGCAAAATGCAATGAGCAACCAATTCATGTTGTA 660
QY 661 ATTGAAGACTTGAAGTCTCACTTACCCGTTGGGTACCCCACTCTCAAGAGCTCTG 720
```

```
DB 661 ATTGAAGACTTGAAGTCTCACTTACCCGTTGGGTACCCCACTCTCAAGAGCTCTG 720
QY 721 AATATTTCTGTTTACCCCGAATATTATTAAGATACATCTGAGAGCTCTGCGCAG 780
DB 721 AATATTTCTGTTTACCCCGAATATTATTAAGATACATCTGAGAGCTCTGCGCAG 780
QY 781 GAGAAAGCAAGATCTGTGACTTCAAGATTCAGATTTTCAAGTGCATTTCAAG 840
DB 781 GAGAAAGCAAGATCTGTGACTTCAAGATTCAGATTTTCAAGTGCATTTCAAG 840
QY 841 GAGTTCAACTTACTAGATGATGATCCATAAACCACTCTGCTGTGATTTGACATT 900
DB 841 GAGTTCAACTTACTAGATGATGATCCATAAACCACTCTGCTGTGATTTGACATT 900
QY 901 TCAATATCAGACTTTTCTATGAGCTGAGATGCTTCAAGGTATCTGCTTAACT 960
DB 901 TCAATATCAGACTTTTCTATGAGCTGAGATGCTTCAAGGTATCTGCTTAACT 960
QY 961 GATTCTGAGGTACAAAGCTTACTCCAAAGACTGACGTTGAAATTAAGAGACCTGC 1020
DB 961 GATTCTGAGGTACAAAGCTTACTCCAAAGACTGACGTTGAAATTAAGAGACCTGC 1020
QY 1021 GTCTTTTGAATAAAGGACAGACAAAGAAAGAGCTTACCTTACCCGACATATA 1080
DB 1021 GTCTTTTGAATAAAGGACAGACAAAGAAAGAGCTTACCTTACCCGACATATA 1080
QY 1081 CTTGCGGAGATTTCTCTCAGTTCAATTTTACCTGTGTCTTGAATCCGACATTTCT 1140
DB 1081 CTTGCGGAGATTTCTCTCAGTTCAATTTTACCTGTGTGTCTTGAATCCGACATTTCT 1140
QY 1141 AAAAAGGATTTTGGAGCCCTGTGACATTAACAGTACAGTGTGTAAGAGGACAG 1200
DB 1141 AAAAAGGATTTTGGAGCCCTGTGACATTAACAGTACAGTGTGTAAGAGGACAG 1200
QY 1201 CTACAGAGCTGTGACATTAACAGGAGGACGAGATTAAGCCGTTGTATGAGATGCC 1260
DB 1201 CTACAGAGCTGTGACATTAACAGGAGGACGAGATTAAGCCGTTGTATGAGATGCC 1260
QY 1261 TGTGCTGCTGTGTGATCTCTCTCGTTCCCTTCTGTCAGCCACACTCAGTCTC 1320
DB 1261 TGTGCTGCTGTGTGATCTCTCTCGTTCCCTTCTGTCAGCCACACTCAGTCTC 1320
QY 1321 CTGCTCGAATCTTCTCTTAAATTTCAACCCAGACATTAAGTGTGACAGCTCAAGTTA 1380
DB 1321 CTGCTCGAATCTTCTCTTAAATTTCAACCCAGACATTAAGTGTGACAGCTCAAGTTA 1380
QY 1381 TTTCAACGAGAAAGCTCAATTTGTCTTCAACATTTGTGAAATTTCTGTCTACGACA 1440
DB 1381 TTTCAACGAGAAAGCTCAATTTGTCTTCAACATTTGTGAAATTTCTGTCTACGACA 1440
QY 1441 ACAAGAGTTCTGCGAAGGAGATATGACAGGCTGAGCTTGTGTGTTGCTTCAAGTT 1500
DB 1441 ACAAGAGTTCTGCGAAGGAGATATGATGACAGGCTGAGCTTGTGTGTTGCTTCAAGTT 1500
QY 1501 CTTGAGCCAAATATCATGATCTCCATGAAACAACCGGAAAGCCCTGCTCTTAAGATA 1560
DB 1501 CTTGAGCCAAATATCATGATCTCCATGAAACAACCGGAAAGCCCTGCTCTTAAGATA 1560
QY 1561 TCCATCTCTCTCGAACAACAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620
DB 1561 TCCATCTCTCTCGAACAACAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620
QY 1621 ATTAATGTGGGTCAAGAACCGGATAGCCCGTTATTTGGTCTTACAACT---AG 1676
DB 1621 ATTAATGTGGGTCAAGAACCGGATAGCCCGTTATTTGGTCTTACAACT---AG 1676
QY 1677 AATCTCAAGAAACAACCGAGATGAAATTTTGAAGCAATGTGTTTGTGGTGC 1736
DB 1677 AATCTCAAGAAACAACCGAGATGAAATTTTGAAGCAATGTGTTTGTGGTGC 1736
QY 1737 AGGCAATTAAGATGAGATTAATCTATTCAGAAAGAGCTCAGACATTTCTTAAGATGAG 1796
```



Db	1741	AGGCATAAGGATNAGGANTATCTANTCAGAAAAGCTCAGACATTTCTTAACGATGGG	1800
Qy	1797	ATCTTAACTCATCTAAAGGTTTCTCTCTCAAGAGATGCTCTGTGGGAGAGAAAGCC	1856
Db	1801	ATCTTAACTCATCTAAAGGTTTCTCTCTCAAGAGATGCTCTGTGGGAGAGAAAGCC	1860
Qy	1857	CCAGCAAAAGTATGTACAAAGCAATCCAGCTTCATGGCCAGCAGTGGCAGAGATCCTC	1916
Db	1861	CCAGCAAAAGTATGTACAAAGCAATCCAGCTTCATGGCCAGCAGTGGCAGAGATCCTC	1920
Qy	1917	CTCCAGAGAAACGGCCATATTTATGTGTGTGGAGTGCAAAGAAATATGGCCAGAGATGTA	1976
Db	1921	CTCCAGAGAAACGGCCATATTTATGTGTGTGGAGTGCAAAGAAATATGGCCAGAGATGTA	1980
Qy	1977	CATGATGCCCTTGTGCAAAATATATAGCAAAAGGTTGAGTTGAAAACTGAAGCATG	2036
Db	1981	CATGATGCCCTTGTGCAAAATATATAGCAAAAGGTTGAGTTGAAAACTGAAGCATG	2040
Qy	2037	AAAAACCTGTGGCACTTTTAAAAAGAAAGAAACGCTACCTTCAGAGATTTTGGTCA	2090
Db	2041	AAAAACCTGTGGCACTTTTAAAAAGAAAGAAACGCTACCTTCAGAGATTTTGGTCA	2094

CC	Result 8
XX	AA58935
ID	AA58935 standard; DNA; 3259 BP.
XX	AA58935;
AC	
XX	
DT	07-NOV-2000 (first entry)
XX	
DE	DNA encoding a human methionine synthase reductase polypeptide.
XX	
KW	Human; methionine synthase reductase; MTRR; cancer;
KW	cardiovascular disease; Down's Syndrome; neural tube defect;
KW	premature coronary artery disease; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Location/Qualifiers
FT	CDS
FT	80..2176
FT	/*tag= a
FT	/product= "methionine synthase reductase"
XX	
PN	W0200042196-A2.
XX	
PD	20-JUL-2000.
XX	
PF	14-JAN-2000; 2000WO-IB000209.
XX	
PR	15-JAN-1999; 99US-00232028.
PR	10-AUG-1999; 99US-00371347.
XX	
PA	(UTMC-) UNIV MCGILL.
XX	
PI	Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;
XX	
DR	WPI; 2000-466131/40.
XX	
XX	P-PSDB; AAB07591.
XX	
PT	Mammalian methionine synthase reductase nucleic acid used for detecting
PT	an increased risk of developing a neural tube defect, Down's Syndrome or
PT	cardiovascular disease in a mammalian embryo or fetus.
XX	
PS	Claim 3; Fig 3; 85pp; English.
XX	
CC	The present sequence encodes a human methionine synthase reductase (MTRR)
CC	polypeptide. Inhibitors of MTRR polypeptide and polynucleotide are used
CC	for treating or preventing cancer, cardiovascular disease, Down's
CC	Syndrome or neural tube defects in a subject. The cardiovascular disease
CC	is premature coronary artery disease. The compounds are detected by
CC	methods which screen for modulators of MTRR biological activity. MTRR
CC	polypeptide or nucleic acid is examined for the presence of a

CC	polymorphism in the parents or the embryo or foetus, and the information
CC	used for detecting an increased risk of an embryo or foetus developing
CC	cancer, cardiovascular disease, Down's Syndrome or neural tube defects
XX	
SQ	Sequence 3259 BP; 944 A; 706 C; 663 G; 946 T; 0 U; 0 Other;
Query Match	80.8%; Score 1691; DB 3; Length 3259;
Best Local Similarity	99.7%; Pred. No. 0;
Matches 2091;	Conservative 0; Mismatches 2; Indels 4; Gaps 1;
QY	1 ATGAGGAGGTTTCGTCTACTATATATGCTACACAGCAGGAGCAAGGCCATCGCAGAA 60
DB	80 ATGAGGAGGTTTCGTCTACTATATGCTACACAGCAGGAGCAAGGCCATCGCAGAA 139
QY	61 GAAATGTGTGAGCAGAGCTGTGTACATGATTTTCTGCAGATCTTCACTGTATTAGTAA 120
DB	140 GAAATGTGTGAGCAGAGCTGTGTACATGATTTTCTGCAGATCTTCACTGTATTAGTAA 199
QY	121 TCCGATTAAGTTAGCCTTAAACCCGAAACAGCTCCTTGTGTGTGTGTTTCTAACACG 180
DB	200 TCCGATTAAGTTAGCCTTAAACCCGAAACAGCTCCTTGTGTGTGTGTTTCTAACACG 259
QY	181 GGCAACCGAGACCCACCCGACACAGCCCGGAGGTTGTAAAGAAATACAGAAACAACA 240
DB	260 GGCAACCGAGACCCACCCGACACAGCCCGGAGGTTGTAAAGAAATACAGAAACAACA 319
QY	241 CTGCGGTTGATTTCTTTGGCTCACCTGCGGTATGGGTTACTGGGTCCTGGTATTCGAA 300
DB	320 CTGCGGTTGATTTCTTTGGCTCACCTGCGGTATGGGTTACTGGGTCCTGGTATTCGAA 379
QY	301 TACACCTTACTTTTGCATATGGGGGGAAGATTAATTGATTAAGACTTTCAGAGCTTGGACC 360
DB	380 TACACCTTACTTTTGCATATGGGGGGAAGATTAATTGATTAAGACTTTCAGAGCTTGGACC 439
QY	361 CGGCATTTCTATGACACTGACACATGACATGACTGTGTATGATTGTAGACTTGTGTGTAG 420
DB	440 CGGCATTTCTATGACACTGACACATGACATGACTGTGTATGATTGTAGACTTGTGTGTAG 499
QY	421 CCGTGAATGTGTGACACTCTGGCCAGCCCTCAGAAAGCAATTTTAGTCAAGCAGAGCAA 480
DB	500 CCGTGAATGTGTGACACTCTGGCCAGCCCTCAGAAAGCAATTTTAGTCAAGCAGAGCAA 559
QY	481 GAGGAGATTAAGTGGCGCACTCCCGGTGGCATCACCTGCAATCTTTGAGGACAGACCTTGTG 540
DB	560 GAGGAGATTAAGTGGCGCACTCCCGGTGGCATCACCTGCAATCTTTGAGGACAGACCTTGTG 619
QY	541 AAGTCAGAGCTGTGACACATTTGATTCACAGTGAGGCTTCAGATTCGATGATTCAGGA 600
DB	620 AAGTCAGAGCTGTGACACATTTGATTCACAGTGAGGCTTCAGATTCGATGATTCAGGA 679
QY	601 AGAAAGGATTCGTAGGTTTGTGAAGCAAAATGCAAGTGAACAGCAACCAATCCAAATGTGTGA 660
DB	680 AGAAAGGATTCGTAGGTTTGTGAAGCAAAATGCAAGTGAACAGCAACCAATCCAAATGTGTGA 739
QY	661 ATTGAAGACTTTGAGTCTCTCACTTACCCTGTTGGTATCCCCACTCTTCAACAAGCTCTCTG 720
DB	740 ATTGAAGACTTTGAGTCTCTCACTTACCCTGTTGGTATCCCCACTCTTCAACAAGCTCTCTG 799
QY	721 AATATTCCTGTGTTTACCCTCCAGAAATTTTACAGGTATCTTGCAGAGGATCTCTTGGCCAG 780
DB	800 AATATTCCTGTGTTTACCCTCCAGAAATTTTACAGGTATCTTGCAGAGGATCTCTTGGCCAG 859
QY	781 GAGGAAAGCCCAAGATCTGTGACTTTCAGACAGATCCAGTTTTCAGAGTCCCAATTTCAAG 840
DB	860 GAGGAAAGCCCAAGATCTGTGACTTTCAGACAGATCCAGTTTTCAGAGTCCCAATTTCAAG 919
QY	841 GCAGTTCAACTTACTACGATGATGCCATTAACCAACTCTGCTGTGTAGAAATTTGCAATT 900
DB	920 GCAGTTCAACTTACTACGATGATGCCATTAACCAACTCTGCTGTGTAGAAATTTGCAATT 979
QY	901 TCAATATCAGACTTTTCTATCAGCTTGGAGATGCTTTCAGCTGATCTGCCCTTAACAGT 960
DB	980 TCAATATCAGACTTTTCTATCAGCTTGGAGATGCTTTCAGCTGATCTGCCCTTAACAGT 1039



QY	961	GATTCTGAGGTACAAAGCTCTCTCCAAAGA	CTGAGAGCTGAAAGTAAAGAAGACATCGC	1020
Db	1040	GATTCTGAGGTACAAAGCTCTCTCCAAAGA	CTGAGAGCTGAAAGTAAAGAAGACATCGC	1039
QY	1021	GTCCCTTTTGAAAATAAAGGCAGACACAAAGAA	AGAGAGCTACCTTACCCAGCATATTA	1080
Db	1100	GTCCCTTTTGAAAATAAAGGCAGACACAAAGAA	AGAGAGCTACCTTACCCAGCATATTA	1159
QY	1081	CTGCGGGGATGTCTCTCCAGTTCAATTTTA	ACTGAGTCTTTGAAATCCAGGCAATTCCT	1140
Db	1160	CTGCGGGGATGTCTCTCCAGTTCAATTTTA	ACTGAGTCTTTGAAATCCAGGCAATTCCT	1219
QY	1141	AAAAAGGCATTTTTCGAGAGCCCTTGATGAC	TATATACAGTGCAGTGTGAAAAGCGCAGG	1200
Db	1220	AAAAAGGCATTTTTCGAGAGCCCTTGATGAC	TATATACAGTGCAGTGTGAAAAGCGCAGG	1279
QY	1201	CTACAGAGAGCTGTGACAGTAAACAAGGCGAG	CCGATTAATAGCCGCTTTGTACAGATATCC	1260
Db	1280	CTACAGAGAGCTGTGACAGTAAACAAGGCGAG	CCGATTAATAGCCGCTTTGTACAGATATCC	1339
QY	1261	TGTGSCCTGTGTGTGGATATCTCTCTGCTTC	CTGCGCAGGACCACTACAGTCTC	1320
Db	1340	TGTGSCCTGTGTGTGGATATCTCTCTGCTTC	CTGCGCAGGACCACTACAGTCTC	1399
QY	1321	CTGCTCGCAACATCTTCCCTAAACTTCAACC	AGACATATTTGTGTGCAAGTCAAGTTTA	1380
Db	1400	CTGCTCGCAACATCTTCCCTAAACTTCAACC	AGACATATTTGTGTGCAAGTCAAGTTTA	1459
QY	1381	TTTTCACCCAGGAAAGCTTCATTTTGTCTTCA	CAACTTGTGGAAATTTCTGTCTACAGCCACA	1440
Db	1460	TTTTCACCCAGGAAAGCTTCATTTTGTCTTCA	CAACTTGTGGAAATTTCTGTCTACAGCCACA	1519
QY	1441	ACAGAGGTTCTGCGGAAAGGAGTATGTACAGG	CTGAGCTGTGTGTGCTTGTGCTTCACTCAGTT	1500
Db	1520	ACAGAGGTTCTGCGGAAAGGAGTATGTACAGG	CTGAGCTGTGTGTGCTTGTGCTTCACTCAGTT	1579
QY	1501	CTTCAGGCCAAACATACATGATATCCATGAA	GACAGCGGAAAGCCCTGTGCTCTTAAGATA	1560
Db	1580	CTTCAGGCCAAACATACATGATATCCATGAA	GACAGCGGAAAGCCCTGTGCTCTTAAGATA	1639
QY	1561	TTCATCTCTCTCGAACAACAATTTCTTTCAC	ATTACCAAGATGACCCCTCAATCCCATC	1620
Db	1640	TTCATCTCTCTCGAACAACAATTTCTTTCAC	ATTACCAAGATGACCCCTCAATCCCATC	1699
QY	1621	ATATATGTGGGTCCAGGAAACCGGATATAGC	CCCGTTTATGTGGGTCTTACACATATAGAGAG	1759
Db	1700	ATATATGTGGGTCCAGGAAACCGGATATAGC	CCCGTTTATGTGGGTCTTACACATATAGAGAG	1759
QY	1677	AAACTCCAGAAACAACCCAGATGAAATTTTGA	GACATGTGTGTTTTTGTGGCTGC	1736
Db	1760	AAACTCCAGAAACAACCCAGATGAAATTTTGA	GACATGTGTGTTTTTGTGGCTGC	1819
QY	1737	AGGCATTAAGATTAAGGATTATCTATTTGAA	AAAGAGCTCACAATTTCTTAAAGATAGG	1796
Db	1820	AGGCATTAAGATTAAGGATTATCTATTTGAA	AAAGAGCTCACAATTTCTTAAAGATAGG	1879
QY	1797	ATCTTAACTCATTAAGGTTTCTTCTTCAAGA	GTCTCTGTGTGGAGAGAGAACCC	1856
Db	1880	ATCTTAACTCATTAAGGTTTCTTCTTCAAGA	GTCTCTGTGTGGAGAGAGAACCC	1939
QY	1857	CCAGCAAAAGTATGTACAGAAACAATCCAGT	TATATGCGAGCGAGAGATCCTC	1916
Db	1940	CCAGCAAAAGTATGTACAGAAACAATCCAGT	TATATGCGAGCGAGAGATCCTC	1999
QY	1917	CTCCAGAGAAACGGCATATTTATGTGTGTGA	GTGCAAGATATATGCGCAAGATGTA	1976
Db	2000	CTCCAGAGAAACGGCATATTTATGTGTGTGA	GTGCAAGATATATGCGCAAGATGTA	2059
QY	1977	CATGATGCCCTTGTGCAAAATATATAGCAAA	AGAGTTGAGATTTGAAAACCTAGAACATAG	2036
Db	2060	CATGATGCCCTTGTGCAAAATATATAGCAAA	AGAGTTGAGATTTGAAAACCTAGAACATAG	2119

Cy	203	AAACCCGCGCAGCTTTAAAAAAGAAGAAAAGCATCTTCAGGATATTGGCATPA	2093
Dd	2120	AAACCCTGGCACCTTTAAAAAGAAAGAAAACGTACTTCAAGATATTGGTCATPA	2176
<hr/>			
RESULT_9			
ID	ADM43216	standard; cDNA; 2091 BP.	
XX	AD43216;		
DT	03-JUN-2004	(first entry)	
DE	Human methionine synthase reductase CDS del 1726-1728 variant.		
XX			
KW	Human; ss; Methionine synthase reductase polypeptide; HsMTRR; cancer;		
KV	cardiovascular disease; neural tube defect; hyperhomocysteinemia;		
KX	chromosome 5p15.2-p15.3; SNP; single nucleotide polymorphism.		
OS	Homo sapiens.		
Key	Location/Qualifiers		
CDS	1..2091		
FT	/tag= a		
FT	/product= "HsMTRRdel1559"		
FT	/partial		
FT	/note= "No stop codon shown"		
FT	replace(66,A)		
FT	/tag= b		
FT	/standard_name= "single_nucleotide_polymorphism"		
FT	replace(110,A)		
FT	/tag= c		
FT	/standard_name= "single_nucleotide_polymorphism"		
FT	replace(1726,TTGT)		
FT	/tag= d		
XX	US2003082676-A1.		
PD	01-MAY-2003.		
PE	10-AUG-1999;	99US-00371347.	
XX			
PR	16-JAN-1998;	98US-0071622P.	
PR	15-JAN-1999;	99US-00232028.	
XX			
PA	(GRAV//) GRAVEL R. A.		
PA	(ROZE//) ROZEN R.		
PA	(LECL//) LECLERC D.		
PA	(WILS//) WILSON A.		
PA	(ROSE//) ROSENBLATT D.		
PI	Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;		
XX			
DR	WPI: 2003-576610/54.		
DR	P-PSDB; ADM43217.		
XX			
PT	New substantially pure nucleic acid encoding a mammalian methionine synthase reductase polypeptide, useful for diagnosing, preventing or treating conditions associated with altered methionine synthase activity, e.g. cancer.		
PT			
XX			
PS	Disclosure; SEQ ID NO 45; 26pp; English.		
XX			
CC	The invention relates to a substantially pure nucleic acid that encodes a mammalian methionine synthase reductase polypeptide, HsMTRR, or that hybridizes at high stringency to a nucleic acid appearing as ADM43208 or CC ADM43209. Also included are a non-human animal where one or both genetic alleles encoding the methionine synthase reductase polypeptide are mutated, an antibody that specifically binds the above methionine synthase reductase polypeptide, a method of detecting the presence of the above polypeptide, a method for detecting sequence variants for CC methionine synthase reductase in a mammal, methods of treating or CC preventing cancer (or cardiovascular disease or neural tube defects) in a		

CC subject, methods of screening for a compound that modulates methionine  
CC synthase reductase biological activity and a method for detecting an  
CC increased risk of developing a neural tube defect in a mammalian embryo  
CC or foetus. The nucleic acid is useful in diagnosing, preventing or  
CC treating conditions associated with altered methionine synthase activity,  
CC such as cancer, cardiovascular disease or neural tube defects, or in  
CC screening for a compound that modulates methionine synthase reductase  
CC biological activity. Naturally occurring variants of the polypeptide are  
CC also associated with hyperhomocysteinemia. The gene for hSMTR is  
CC located on chromosome 5p15.2-p15.3. The present sequence is the coding  
CC sequence of a variant human hSMTR cDNA.  
XX

Sequence 2091 BP; 591 A; 489 C; 480 G; 531 T; 0 U; 0 Other;

Query Match 80.1%; Score 1677; DB 11; Length 2091;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1677; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAGAGGTTTCTGTTACTATATGCTACACAGAGGACAGGCAAGGCCATCGCAGAA 60  
DB 1 ATGAGAGGTTTCTGTTACTATATGCTACACAGAGGACAGGCAAGGCCATCGCAGAA 60  
QY 61 GAAATGTGAGCAAGCTGTGTACATGATGATTTCTGCAGATCTTCACGTATTAAGTAA 120  
DB 61 GAAATGTGAGCAAGCTGTGTACATGATGATTTCTGCAGATCTTCACGTATTAAGTAA 120  
QY 121 TCCGATATAGTATGACCTAAACCGAAACGCTCTTGTGTGTGTGTGTTTCTACACG 180  
DB 121 TCCGATATAGTATGACCTAAACCGAAACGCTCTTGTGTGTGTGTGTTTCTACACG 180  
QY 181 GGCACCGGAGACCCACCCGACACAGCCCGCAAGTTTGTAAAGAAATACGAACCAACA 240  
DB 181 GGCACCGGAGACCCACCCGACACAGCCCGCAAGTTTGTAAAGAAATACGAACCAACA 240  
QY 241 CTGCGGTTGATTTCTTGTCTCACTGCGGTATGGGTTTCTGGGCTCGGTGATTCAGAA 300  
DB 241 CTGCGGTTGATTTCTTGTCTCACTGCGGTATGGGTTTCTGGGCTCGGTGATTCAGAA 300  
QY 301 TACACCTACTTTTGCATATGGGGGAGAAATATTGATTAACGACTTCAAGAGCTTGAAGCC 360  
DB 301 TACACCTACTTTTGCATATGGGGGAGAAATATTGATTAACGACTTCAAGAGCTTGAAGCC 360  
QY 361 CGGCATTTCTATGACATGACATGACATGACATGACATGACATGACATGACATGACATG 420  
DB 361 CGGCATTTCTATGACATGACATGACATGACATGACATGACATGACATGACATGACATG 420  
QY 421 CCGTGAATGCTGGAATCTGCGACAGCCCTCAGAAAGCAATTTAGTCAAGCAGAGGACAA 480  
DB 421 CCGTGAATGCTGGAATCTGCGACAGCCCTCAGAAAGCAATTTAGTCAAGCAGAGGACAA 480  
QY 481 GAGGAGATTAAGTGGCGCACTCCGGTGGCATCACTGCACTCTTGAAGGACAGACCTTGTG 540  
DB 481 GAGGAGATTAAGTGGCGCACTCCGGTGGCATCACTGCACTCTTGAAGGACAGACCTTGTG 540  
QY 541 AAGTCAGAGCTGCTACATGATGATCTCAAGTGAAGTCTTGAAGTCAAGTCAAGTCAAG 600  
DB 541 AAGTCAGAGCTGCTACATGATGATCTCAAGTGAAGTCTTGAAGTCAAGTCAAGTCAAG 600  
QY 601 AGAAGGATTTCTGAGTTTGAAGCAAAATGACGTGAACAGCAACCAATCCAAATGTTGA 660  
DB 601 AGAAGGATTTCTGAGTTTGAAGCAAAATGACGTGAACAGCAACCAATCCAAATGTTGA 660  
QY 661 ATTGAAGACTTTGAGTCTTCACTTACCGTTCGTTACCCCACTCTCAAGGCTCTCTG 720  
DB 661 ATTGAAGACTTTGAGTCTTCACTTACCGTTCGTTACCCCACTCTCAAGGCTCTCTG 720  
QY 721 AATATTCCTGTTTACCCCAAGATATTTACAGTATATCTGCAAGGAGTCTCTTGGCAG 780  
DB 721 AATATTCCTGTTTACCCCAAGATATTTACAGTATATCTGCAAGGAGTCTCTTGGCAG 780  
QY 781 GAGGAAAGCCAAAGTATCTGATCTTCAAGATCAGTTTCAAGTCCAAATTTCAAG 840  
DB 781 GAGGAAAGCCAAAGTATCTGATCTTCAAGATCAGTTTCAAGTCCAAATTTCAAG 840

QY 841 GCAATTCACTTACTACGATGATGCTATTAACCACTCTGCTGTATGATTTGACATT 900  
DB 841 GCAATTCACTTACTACGATGATGCTATTAACCACTCTGCTGTATGATTTGACATT 900  
QY 901 TCAATATACAGCTTTTCTATGAGCTGTGAGATGCTTGAAGGTATGCTGCTTACAGT 960  
DB 901 TCAATATACAGCTTTTCTATGAGCTGTGAGATGCTTGAAGGTATGCTGCTTACAGT 960  
QY 961 GATTCGAGTACAAAGCTTACTCCAAAGATGACGCTTGAAGATTAAGAGACAGCTGC 1020  
DB 961 GATTCGAGTACAAAGCTTACTCCAAAGATGACGCTTGAAGATTAAGAGAGCACTGC 1020  
QY 1021 GTTCCTTTGAAAATTAAGGACACACAAAGAAAGAGCTTACCTTACCCAGCATTA 1080  
DB 1021 GTTCCTTTGAAAATTAAGGACACACAAAGAAAGAGCTTACCTTACCCAGCATTA 1080  
QY 1081 CCGCGGGAGTGTCTCTCAGATTCAATTTTACCTGCTGATCGGAATCCGAGCAATTCT 1140  
DB 1081 CCGCGGGAGTGTCTCTCAGATTCAATTTTACCTGCTGATCGGAATCCGAGCAATTCT 1140  
QY 1141 AAAAAGCAATTTTGGAGCCCTTGTGACTATACAGTGAAGTGTGAAAAAGCCAGG 1200  
DB 1141 AAAAAGCAATTTTGGAGCCCTTGTGACTATACAGTGAAGTGTGAAAAAGCCAGG 1200  
QY 1201 CTACAGAGCTGTGACGTAACAAAGGAGGAGCCGATTAATAGCCGTTTGTAGAGATGCC 1260  
DB 1201 CTACAGAGCTGTGACGTAACAAAGGAGGAGCCGATTAATAGCCGTTTGTAGAGATGCC 1260  
QY 1261 TGTGCTGCTGTGTGATCTCTCCCTCGCTGCTTCCCTTCCGACAGCAACGCTCAATCTC 1320  
DB 1261 TGTGCTGCTGTGTGATCTCTCCCTCGCTGCTTCCCTTCCGACAGCAACGCTCAATCTC 1320  
QY 1321 CTGCTGGAACATCTTCTTAACTTCAACCCAGACATATTCGTGCAAGCTCAAGTTTA 1380  
DB 1321 CTGCTGGAACATCTTCTTAACTTCAACCCAGACATATTCGTGCAAGCTCAAGTTTA 1380  
QY 1381 TTTCAACCCAGAAAGCTCCATTTTGTCTTCAACATTTGTGAATTTCTGTCTACGCA 1440  
DB 1381 TTTCAACCCAGAAAGCTCCATTTTGTCTTCAACATTTGTGAATTTCTGTCTACGCA 1440  
QY 1441 ACAGAGGTTCTGCGGAAGGAGATGATGACAGGCTGAGCTTGTGTGCTTCAAGTT 1500  
DB 1441 ACAGAGGTTCTGCGGAAGGAGATGATGACAGGCTGAGCTTGTGTGCTTCAAGTT 1500  
QY 1501 CTTCAAGCCAAACATATCATGATCCCATGAAGACAGGAGGAAAGCCCTGGCTCTTAAGATA 1560  
DB 1501 CTTCAAGCCAAACATATCATGATCCCATGAAGACAGGAGGAAAGCCCTGGCTCTTAAGATA 1560  
QY 1561 TCCATCTCTCTCGAACAACAAATTTTTCACCTTACAGATGACCCCTCAATCCCATC 1620  
DB 1561 TCCATCTCTCTCGAACAACAAATTTTTCACCTTACAGATGACCCCTCAATCCCATC 1620  
QY 1621 ATTAATGTGGGTCCAGGAACCGGACATAGCCCGTTATTTGGGTTCTTCAACAATAGA 1677  
DB 1621 ATTAATGTGGGTCCAGGAACCGGACATAGCCCGTTATTTGGGTTCTTCAACAATAGA 1677  
RESULT 10  
AD087538 standard; cDNA; 3270 BP.  
AD087538;  
07-OCT-2004 (first entry)  
Human tumour-associated antigenic target (TAT) cDNA sequence #4416.  
human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;  
cancer; cell proliferative disorder; gene; ss.  
Homo sapiens.  
XX



QY 1201 CTACAGAGCTGTGCTAGTAAACAAGGGGACCGGATTAATAGCCGCTTTGTACAGATGCC 1260  
DB 1312 CTACAGAGCTGTGCTAGTAAACAAGGGGACCGGATTAATAGCCGCTTTGTACAGATGCC 1371  
QY 1261 TGGGCGCTGTGTGATCTCTCTGCTTCCCTTCTTGCCAGCCACACACAGCTC 1320  
DB 1372 TGGGCGCTGTGTGATCTCTCTGCTTCCCTTCTTGCCAGCCACACACAGCTC 1431  
QY 1321 CTGCTGAAATCTTCTTAACTTCAACCCAGACATATTCGTGCAAGCTCAAGTTTA 1380  
DB 1432 CTGCTGAAATCTTCTTAACTTCAACCCAGACATATTCGTGCAAGCTCAAGTTTA 1491  
QY 1381 TTTCACCCAGAAAAGCTCCATTTTGTCTTCAACATGTGGAATTTCTGTCTACTGCCACA 1440  
DB 1492 TTTCACCCAGAAAAGCTCCATTTTGTCTTCAACATGTGGAATTTCTGTCTACTGCCACA 1551  
QY 1441 ACAGAGGTTCTGGGGAAGGAGTATGTAAGGCGGTGCTGCTTGTGTGCTCAAGTT 1500  
DB 1552 ACAGAGGTTCTGGGGAAGGAGTATGTAAGGCGGTGCTGCTTGTGTGCTCAAGTT 1611  
QY 1501 CTTCAGCCAAACATACATGATCCCATGAAGAAGCCGCGGAAAGCCCTGCTCTTAAGATA 1560  
DB 1612 CTTCAGCCAAACATACATGATCCCATGAAGAAGCCGCGGAAAGCCCTGCTCTTAAGATA 1671  
QY 1561 TCCTATCTCTCTGGAACAACAATTTCTTCCACTTACCAATGACCCCTCAATCCCATC 1620  
DB 1672 TCCTATCTCTCTGGAACAACAATTTCTTCCACTTACCAATGACCCCTCAATCCCATC 1731  
QY 1621 ATATAGTGTGGTCCAGGACCGGATATGCCCCGTTTATTTGGGTTTCTTCAACAAT---AG 1676  
DB 1732 ATATAGTGTGGTCCAGGACCGGATATGCCCCGTTTATTTGGGTTTCTTCAACAAGAG 1791  
QY 1677 AAATCCCAAGAACACACCCAGATGGAATTTTGGAGCAATGTGTTTGTGTTTGGCTGC 1736  
DB 1792 AAATCCCAAGAACACACCCAGATGGAATTTTGGAGCAATGTGTTTGTGTTTGGCTGC 1851  
QY 1737 AGGCTATAGATAGGATTTTCTATTTGAGAAAAGCTCGACATTTCTTTAAGCATGGG 1796  
DB 1852 AGGCTATAGATAGGATTTTCTATTTGAGAAAAGCTCGACATTTCTTTAAGCATGGG 1911  
QY 1797 ATCTTAACTGATCTTAAAGTTTCTCTCTCAAGAGATCTCTGTGTTGGGAGAGGAAGCC 1856  
DB 1912 ATCTTAACTGATCTTAAAGTTTCTCTCTCAAGAGATCTCTGTGTTGGGAGAGGAAGCC 1971  
QY 1857 CCAGCAAGATATGTAACAACAATCAGCTTCATGCGCAGCAGGATGCGAGATCTCTC 1916  
DB 1972 CCAGCAAGATATGTAACAACAATCAGCTTCATGCGCAGCAGGATGCGAGATCTCTC 2031  
QY 1917 CTCCAGGAGAACGGCCATTTATGTGTGTGAGATGCAAAAGATATGCGCAAGATGTA 1976  
DB 2032 CTCCAGGAGAACGGCCATTTATGTGTGTGAGATGCAAAAGATATGCGCAAGATGTA 2091  
QY 1977 CATGATCCCTTGTGCAAAATTAATAAGCAAGAGTTGAGTGAATAAATCTAGAAGCAATG 2036  
DB 2092 CATGATCCCTTGTGCAAAATTAATAAGCAAGAGTTGAGTGAATAAATCTAGAAGCAATG 2151  
QY 2037 AAAACCTGCGCACTTTAAAGAAAGAAAACGCTACCTTCAGGATATTTGGTCATTA 2093  
DB 2152 AAAACCTGCGCACTTTAAAGAAAGAAAACGCTACCTTCAGGATATTTGGTCATTA 2208

RESULT 11  
AAA58977  
ID AAA58977 standard; DNA; 3256 BP.  
XX  
XX  
AC AAA58977;  
XX  
DT 07-NOV-2000 (first entry)  
XX  
DE A human methionine synthase reductase DNA sequence with polymorphism.  
XX  
KM Human; methionine synthase reductase; MTRR; cancer;  
KM Cardiovascular disease; Down's Syndrome; neural tube defect;

KW premature coronary artery disease; ss.  
XX  
OS Homo sapiens.  
XX  
PN MO200042196-A2.  
XX  
PD 20-JUL-2000.  
XX  
XX  
PF 14-JAN-2000; 2000MO-IB000209.  
XX  
PR 15-JAN-1999; 99US-00232028.  
PR 10-AUG-1999; 99US-00371347.  
XX  
PA (UVMC-) UNITV MCGILL.  
PI Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;  
XX  
XX WPI; 2000-466131/40.  
XX  
XX  
XX Mammalian methionine synthase reductase nucleic acid used for detecting  
PT an increased risk of developing a neural tube defect, Down's Syndrome or  
PT cardiovascular disease in a mammalian embryo or fetus.  
XX  
XX  
PS Claim 8; Page; 85pp; English.  
XX  
XX The present sequence represents a human methionine synthase reductase  
CC (MTRR) DNA sequence, with a polymorphism comprising of a deletion of  
CC nucleotides 1726-1728. Inhibitors of MTRR polypeptide and polynucleotide  
CC are used for treating or preventing cancer, cardiovascular disease,  
CC Down's Syndrome or neural tube defects in a subject. The cardiovascular  
CC disease is premature coronary artery disease. The compounds are detected  
CC by methods which screen for modulators of MTRR biological activity. MTRR  
CC polypeptide or nucleic acid is examined for the presence of a  
CC polymorphism in the parents or the embryo or foetus, and the information  
CC used for detecting an increased risk of an embryo or foetus developing  
CC cancer, cardiovascular disease, Down's Syndrome or neural tube defects.  
CC note: the present sequence does not appear in the specification; it was  
CC created using information provided  
XX  
SQ Sequence 3256 BP; 943 A; 705 C; 662 G; 946 T; 0 U; 0 Other;

Query Match 76.2%; Score 1595; DB 3; Length 3256;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1645; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGAGGTTTCTGTACTATATGCTACACAGCAGGGAAGGCAAGCCATCGCAGAA 60  
DB 80 ATGAGAGGTTTCTGTACTATATGCTACACAGCAGGGAAGGCAAGCCATCGCAGAA 139  
QY 61 GAAATGTGAGCAAGCTGTGTACATGATTTTCTGCAAGATCTTCACTGTATTAAGTAA 120  
DB 140 GAAATGTGAGCAAGCTGTGTACATGATTTTCTGCAAGATCTTCACTGTATTAAGTAA 199  
QY 121 TCCGATTAAGTAACTTAAACCGGAAGAGCTCTTGTGTGTTGTTTCTACACAG 180  
DB 200 TCCGATTAAGTAACTTAAACCGGAAGAGCTCTTGTGTGTTGTTTCTACACAG 259  
QY 181 GGCACCGAGAACCCACCCGACAGCAGCCGCAAGTTTGTAAAGAAATACAGAACCAACA 240  
DB 260 GGCACCGAGAACCCACCCGACAGCAGCCGCAAGTTTGTAAAGAAATACAGAACCAACA 319  
QY 241 CTGCGGTTGATTTCTTGTCTACCTGCGGTATGAGTTACTGGGTCCTGGTATTCAGAA 300  
DB 320 CTGCGGTTGATTTCTTGTCTACCTGCGGTATGAGTTACTGGGTCCTGGTATTCAGAA 379  
QY 301 TACACTACTTTTGGCAATGGGGGGAAGATTAATGATTAACGACTTCAAGGCTTGAGCC 360  
DB 380 TACACTACTTTTGGCAATGGGGGGAAGATTAATGATTAACGACTTCAAGGCTTGAGCC 439  
QY 361 CGGCATTTCTATGACACTGACATGACATGACTGTGTAGGTTTAAGAACTTGTGTTGAG 420  
DB 440 CGGCATTTCTATGACACTGACATGACATGACTGTGTAGGTTTAAGAACTTGTGTTGAG 499

```

QY 421 CCGTGGATGTCGACCTGCGCCAGAGCCCTCAGAAAGCATTTTAAAGTCAAGCAGAGAGACA 480
DB 500 CCGTGGATGTCGACCTGCGCCAGAGCCCTCAGAAAGCATTTTAAAGTCAAGCAGAGAGACA 559
QY 481 GAGGAGATTAAGTGGGCACTCCCGGTGGGATCACTGCACTCTTGAGGAGACAGCTTTGG 540
DB 560 GAGGAGATTAAGTGGGCACTCCCGGTGGGATCACTGCACTCTTGAGGAGACAGCTTTGG 619
QY 541 AAGTCAGAGCTGTACACATTTGAATCTCAAGTCAGAGCTTCTGAGATTCGATGATTCAGGA 600
DB 620 AAGTCAGAGCTGTACACATTTGAATCTCAAGTCAGAGCTTCTGAGATTCGATGATTCAGGA 679
QY 601 AGAAAGATTCGAGGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATCCATGTTGTA 660
DB 680 AGAAAGATTCGAGGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATCCATGTTGTA 739
QY 661 ATTGAAGACTTGAAGTCTCACTTACCCGTTGGGTACCCCACTCTCAAGCTCTCTG 720
DB 740 ATTGAAGACTTGAAGTCTCACTTACCCGTTGGGTACCCCACTCTCAAGCTCTCTG 799
QY 721 AATATTCCTGTTTACCCCAAGATTTTACAGGTACATCTGAGGAGTCTCTGGCCAG 780
DB 800 AATATTCCTGTTTACCCCAAGATTTTACAGGTACATCTGAGGAGTCTCTGGCCAG 859
QY 781 GAGAAAGCCAAATATCTGATCTTGAAGATCCAGTCTTCAAGTCCAAATTTCAAG 840
DB 860 GAGAAAGCCAAATATCTGATCTTGAAGATCCAGTCTTCAAGTCTTCAAGTCTTCAAG 919
QY 841 GAGATTCAACTTACAGAAATGATGCAATTAACCACTCTGCTGATGATTTGA 900
DB 920 GAGATTCAACTTACAGAAATGATGCAATTAACCACTCTGCTGATGATTTGA 979
QY 901 TCATAATACGACTTTCTTATCAGCCTGAGATGCTTCAAGCTGATCTGCTTACAT 960
DB 980 TCATAATACGACTTTCTTATCAGCCTGAGATGCTTCAAGCTGATCTGCTTACAT 1039
QY 961 GATTCTGAGGTACAAAGCTTACCAAGACTGCACTTGAAGATTAAGAGAGAGACTGC 1020
DB 1040 GATTCTGAGGTACAAAGCTTACCAAGACTGCACTTGAAGATTAAGAGAGAGACTGC 1099
QY 1021 GTCTCTTTGAAATTAAGGAGACACAAAGAAAGAGAGCTTACCTTACCCAGCATATA 1080
DB 1100 GTCTCTTTGAAATTAAGGAGACACAAAGAAAGAGAGCTTACCTTACCCAGCATATA 1159
QY 1081 CCGCGGAGATGTTCTCTCCAGATTCATTTTAACTGCTGATCTTGAATCCGAGCAATTTCT 1140
DB 1160 CCGCGGAGATGTTCTCTCCAGATTCATTTTAACTGCTGATCTTGAATCCGAGCAATTTCT 1219
QY 1141 AAAAAGGCAATTTTGGAGCCCTTGTGACTATACAGTACAGTGTGCTGAAAAGCCAGG 1200
DB 1220 AAAAAGGCAATTTTGGAGCCCTTGTGACTATACAGTGTGCTGAAAAGCCAGG 1279
QY 1201 CTACAGAGCTGTGAGTAAACAAAGGGGAGCCGATTAATAGCCGCTTGTATCAGAGATGCC 1260
DB 1280 CTACAGAGCTGTGAGTAAACAAAGGGGAGCCGATTAATAGCCGCTTGTATCAGAGATGCC 1339
QY 1261 TGTGCTGCTGTTGGATCTCTCTGCTGCTTCCCTTCTTCCAGAGCACCAGTCACTC 1320
DB 1340 TGTGCTGCTGTTGGATCTCTCTGCTGCTTCCCTTCTTCCAGAGCACCAGTCACTC 1399
QY 1321 CTGCTGGAACATTTCTTAACCTTCAACCCAGACCAATTCGATGAGTCAAGTTTA 1380
DB 1400 CTGCTGGAACATTTCTTAACCTTCAACCCAGACCAATTCGATGAGTCAAGTTTA 1459
QY 1381 TTTCAACCCAGAAAGCTCCATTTTGTCTTCAACATTTGTGATTTTGTCTACTGACACA 1440
DB 1460 TTTCAACCCAGAAAGCTCCATTTTGTCTTCAACATTTGTGATTTTGTCTACTGACACA 1519
QY 1441 ACAAGAGTTCTGGGAAAGGAGATGATACAGGCTGGCTGCTTGTGTTGCTTCAGTT 1500
DB 1520 ACAAGAGTTCTGGGAAAGGAGATGATACAGGCTGGCTGCTTGTGTTGCTTCAGTT 1579
QY 1501 CTTCAAGCAAAACATGATGATCCCATGAAAGACAGGGGAAAAGCCCTGGCTCTTAAGATA 1560

```

```

DB 1580 CTTACAGCCAAACATGATGATCCCATGAAGACAGGGGAAAAGCCCTGCTCTTAAGATA 1639
QY 1561 TCCATCTCTCTCGAACAACAATTTCTTCACTTACAGATGAGACCCCTCAATCCCATC 1620
DB 1640 TCCATCTCTCTCGAACAACAATTTCTTCACTTACAGATGAGACCCCTCAATCCCATC 1699
QY 1621 ATAAATGTTGGTCCAGAAACCCGGCAT 1646
DB 1700 ATAAATGTTGGTCCAGAAACCCGGCAT 1725

RESULT 12
AAA58976
ID AAA58976 standard; DNA; 3255 BP.
XX
AC AAA58976;
XX
DT 07-NOV-2000 (first entry)
XX
DE A human methionine synthase reductase DNA sequence with polymorphism.
XX
KW Human; methionine synthase reductase; MTRR; cancer;
KW cardiovascular disease; Down's Syndrome; neural tube defect;
KW premature coronary artery disease; ss.
XX
OS Homo sapiens.
XX
PN WO200042196-A2.
XX
PD 20-JUL-2000.
XX
PF 14-JAN-2000; 2000MO-IB000209.
XX
PR 15-JAN-1999; 99US-00232028.
PR 10-AUG-1999; 99US-00371347.
XX
PA (UWMC-) UNIV MCGILL.
XX
PI Gravel RA, Rozen R, Leclerc D, Wilson A, Rosenblatt D;
XX
DR MPI; 2000-466131/40.
XX
PT Mammalian methionine synthase reductase nucleic acid used for detecting
PT an increased risk of developing a neural tube defect, Down's Syndrome or
PT cardiovascular disease in a mammalian embryo or fetus.
XX
PS Claim 7; Page; 85pp; English.
XX
CC The present sequence represents a human methionine synthase reductase
CC (MTRR) DNA sequence, with a polymorphism comprising of a deletion of
CC nucleotides 1675-1678. Inhibitors of MTRR polypeptide and polynucleotide
CC are used for treating or preventing cancer, cardiovascular disease,
CC Down's Syndrome or neural tube defects in a subject. The cardiovascular
CC disease is premature coronary artery disease. The compounds are detected
CC by methods which screen for modulators of MTRR biological activity. MTRR
CC polypeptide or nucleic acid is examined for the presence of a
CC polymorphism in the parents or the embryo or foetus, and the information
CC used for detecting an increased risk of an embryo or foetus developing
CC cancer, cardiovascular disease, Down's Syndrome or neural tube defects.
CC note: the present sequence does not appear in the specification; it was
CC created using information provided
XX
SQ Sequence 3255 BP; 942 A; 704 C; 663 G; 946 T; 0 U; 0 Other;

Query Match 73.8%; Score 1544; DB 3; Length 3255;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1594; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGAGGTTTCTGTTACTATATGCTTACACAGCAGGAGCAAGGCAATCCGAGAA 60
DB 80 ATGAGAGGTTTCTGTTACTATATGCTTACACAGCAGGAGCAAGGCAATCCGAGAA 139

```





in gene mapping.

Claim 1; Page; 190pp; English.

The invention relates to novel diagnostic and therapeutic polynucleotides selected from one of the 2722 sequences defined in the specification. A polynucleotide of the invention may have a use in gene therapy. The human diagnostic and therapeutic polynucleotides (dthnp) or polypeptides may be used to diagnose a particular condition, disease or disorder associated with human molecules, e.g. cell proliferative disorders, autoimmune/inflammatory disorder, developmental disorder, endocrine disorder, neurological disorders, gastrointestinal disorders, or infections caused by virus, bacteria, fungi or parasite. The dthnp molecules may also be used in genetic mapping, in identifying individuals from minute biological samples, in detecting single nucleotide polymorphisms, as molecular weight markers, and for somatic or germline gene therapy. Note: The sequence data for this patent is not represented in the printed specification, but was obtained in electronic format directly from WIPO at [www.wipo.int/pct/en/sequences/listing.htm](http://www.wipo.int/pct/en/sequences/listing.htm)

Sequence 3189 BP; 916 A; 679 C; 665 G; 929 T; 0 U; 0 Other;

Query Match 45.7%; Score 956; DB 13; Length 3189;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1056; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

OY 1 ATGAGAGAGTTCTGTACTATATGCTACACAGAGGACAGCAAGGCCATCGCAGAA 60
DB 112 ATGAGAGAGTTCTGTACTATATGCTACACAGAGGACAGCAAGGCCATCGCAGAA 171
OY 61 GAAATGTGAGCAAGCTGTGATGATGATTTCTGAGATCTTCACTGATATTAGTAA 120
DB 172 GAAATGTGAGCAAGCTGTGATGATGATTTCTGAGATCTTCACTGATATTAGTAA 231
OY 121 TCCGATATGATGACCTAAACCCGAAACAGCTCTCTTGTGTGTGTGTTCTACACG 180
DB 232 TCCGATATGATGACCTAAACCCGAAACAGCTCTCTTGTGTGTGTGTTCTACACG 291
OY 181 GGCACCGGAGACCCACCCGACACAGCCGAGTTGTTAAGAAATACAGAACCAACA 240
DB 292 GGCACCGGAGACCCACCCGACACAGCCGAGTTGTTAAGAAATACAGAACCAACA 351
OY 241 CCGCGGTTGATTTCTTGTCTACCTGCGGTATGAGTTCTGCGTCTCGGTATTCAGAA 300
DB 352 CCGCGGTTGATTTCTTGTCTACCTGCGGTATGAGTTCTGCGTCTCGGTATTCAGAA 411
OY 301 TACACTACTTTTGCATGCGGGGAGATATGATTAACGACTTCAAGAGCTTGGAGCC 360
DB 412 TACACTACTTTTGCATGCGGGGAGATATGATTAACGACTTCAAGAGCTTGGAGCC 471
OY 361 CCGCATTTCTATGACACTGACATGCAATGATGATGTTGATTAAGCTTGTGTTGAG 420
DB 472 CCGCATTTCTATGACACTGACATGCAATGATGATGTTGATTAAGCTTGTGTTGAG 531
OY 421 CCGTGATATGCGGATCTGCGCCAGCCCTCAAGAAACATTTATGATCAAGAGAGCA 480
DB 532 CCGTGATATGCGGATCTGCGCCAGCCCTCAAGAAACATTTATGATCAAGAGAGCA 591
OY 481 GAGAGATATAGTGGCCACTCCCGGTGGATCAGCTGACCTCTTGGAGACAGACTTTG 540
DB 592 GAGAGATATAGTGGCCACTCCCGGTGGATCAGCTGACCTCTTGGAGACAGACTTTG 651
OY 541 AAGTCAGAGCTGTACACATTTGATTTCAAGTCGAGCTTCTGAGATTCTGATTCAGGA 600
DB 652 AAGTCAGAGCTGTACACATTTGATTTCAAGTCGAGCTTCTGAGATTCTGATTCAGGA 711
OY 601 AGAAGAGATTTGAGGTTTGAAGCAAAATGACGTAAACAGCAACCAATCCATGTTGA 660
DB 712 AGAAGAGATTTGAGGTTTGAAGCAAAATGACGTAAACAGCAACCAATCCATGTTGA 771
OY 661 ATTGAAGACTTGAAGTCTCACTTACCCGTTCCGTAACCCCACTTCAAGGCTTCTG 720

```

```

DB 772 ATTGAAGACTTGAAGTCTCACTTACCCGTTCCGTAACCCCACTTCAAGGCTTCTG 831
OY 721 AATATTCCTGTTTACCCCGAATATTTACAGTACATCTGACAGAGTCTTGGCCAG 780
DB 832 AATATTCCTGTTTACCCCGAATATTTACAGTACATCTGACAGAGTCTTGGCCAG 891
OY 781 GAGAAAGCCAGATATCTGTGACTTCAGAGATCCAGTTTTCAGATGCAATTTGAA 840
DB 892 GAGAAAGCCAGATATCTGTGACTTCAGAGATCCAGTTTTCAGATGCAATTTGAA 951
OY 841 GCAATTCACCTTACTACGATATGATGCAATTAACCACTCTGCTGTAGATTTGACAT 900
DB 952 GCAATTCACCTTACTACGATATGATGCAATTAACCACTCTGCTGTAGATTTGACAT 1011
OY 901 TCAATATGAGACTTTTCCATGAGCTGAGATGCTTCCAGGTATCTGCCATACAT 960
DB 1012 TCAATATGAGACTTTTCCATGAGCTGAGATGCTTCCAGGTATCTGCCATACAT 1071
OY 961 GATTCGAGATCAAAAGCTTACTCCAAAGACTGACCTTGAAGTAAAGAGACATGC 1020
DB 1072 GATTCGAGATCAAAAGCTTACTCCAAAGACTGACCTTGAAGTAAAGAGACATGC 1131
OY 1021 GTCTTTTGAATAAAGGACACACAAAGAAAGG 1058
DB 1132 GTCTTTTGAATAAAGGACACACAAAGAAAGG 1169

```

```

RESULT 14
ADQ39029
ID ADQ39029 standard; DNA; 3256 BP.
XX
AC ADQ39029;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human SNP containing myocardial infarction-associated gene, SEQ ID 692.
XX
KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;
XX
KW cardiant; gene therapy; human; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2004058052-A2.
XX
PD 15-UTL-2004.
XX
PF 22-DEC-2003; 2003WO-US040978.
XX
PR 20-DEC-2002; 2002US-0434778P.
XX
PR 10-MAR-2003; 2003US-0453135P.
XX
PR 30-APR-2003; 2003US-0466122P.
XX
PR 23-SEP-2003; 2003US-0504955P.
XX
PA (APPL-) APPLERA CORP.
XX
PI Cargill M, Devlin JJ, Iakubova O;
XX
DR WPI; 2004-533949/51.
XX
DR P-PSDB; ADQ39857.
XX
PT Identifying an individual who has an altered risk for developing
PT myocardial infarction by detecting a single nucleotide polymorphism in
PT the individual's nucleic acids.
XX
PS Claim 7; SEQ ID NO 692; 145bp; English.
XX
CC The invention relates to a novel method for identifying an individual who
CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least

```







```

Qy 1327 GAACATCTCTTAACCTTCAACCCAGACCAATTCGTGSCAAGCTCAAGTTATTTCAC 1386
    |||||
Db 1438 GAACATCTCTCTTAACCTTCAACCCAGACCAATTCGTGSCAAGCTCAAGTTATTTCAC 1497
    |||||
Qy 1387 CCAGGAAAGCTCATTGTTGTTCTTCAACATTTGTGGAATTTCTGTCTACTGCGACAAGAG 1446
    |||||
Db 1498 CCAGGAAAGCTCATTGTTGTTCTTCAACATTTGTGGAATTTCTGTCTACTGCGACAAGAG 1557
    |||||
Qy 1447 GTTCTGCGAAGGAGATGATGACAGGCTGCGCTTGTGTGTTGCTTCAGTCTTCAG 1506
    |||||
Db 1558 GTTCTGCGAAGGAGATGATGACAGGCTGCGCTTGTGTGTTGCTTCAGTCTTCAG 1617
    |||||
Qy 1507 CCAAACATACATGATCCCATGAAAGACGCGGAAAGCCCTGGCTCTTAAGATATCCATC 1566
    |||||
Db 1618 CCAAACATACATGATCCCATGAAAGACGCGGAAAGCCCTGGCTCTTAAGATATCCATC 1677
    |||||
Qy 1567 TCTCTGGAACAACAATTCCTTCCACTTACAGATGACCCCTCAATCCCATCATPATG 1626
    |||||
Db 1678 TCTCTGGAACAACAATTCCTTCCACTTACAGATGACCCCTCAATCCCATCATPATG 1737
    |||||
Qy 1627 GTGGGTCAGAGAACCGGCATAGCCCGGTTATT 1659
    |||||
Db 1738 GTGGGTCAGAGAACCGGCATAGCCCGGTTATT 1770
    |||||

```

Search completed: August 27, 2005, 01:18:48  
 Job time : 739.091 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 23:30:25 ; Search time 235.308 Seconds  
(without alignments)  
14554.251 Million cell updates/sec

Title: US-09-371-347a-47

Perfect score: 2093  
Sequence: 1 atgagagaggttcgttact.....ttcagatattgtgcataa 2093

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 81813359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq:\*
- 2: /cgn2\_6/prodata/1/ina/5A\_COMB.seq:\*
- 3: /cgn2\_6/prodata/1/ina/5A\_COMB.seq:\*
- 4: /cgn2\_6/prodata/1/ina/5B\_COMB.seq:\*
- 5: /cgn2\_6/prodata/1/ina/5B\_COMB.seq:\*
- 6: /cgn2\_6/prodata/1/ina/Backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1793	85.7	3259	3	US-09-318-448-23 Sequence 23, Appl
2	1640	78.4	3242	4	US-09-949-016-4215 Sequence 4215, Ap
3	386	18.4	390	3	US-08-905-223-71 Sequence 71, Appl
4	330	15.8	601	4	US-09-949-016-150019 Sequence 150019, A
5	279	13.3	601	4	US-09-949-016-150020 Sequence 150020, A
6	189	9.0	601	4	US-09-949-016-150037 Sequence 150037, A
7	158	7.5	2475	4	US-09-566-921-88 Sequence 88, Appl
8	155	7.4	601	4	US-09-949-016-150030 Sequence 150030, A
9	145	6.9	601	4	US-09-949-016-150031 Sequence 150031, A
10	137	6.5	601	4	US-09-949-016-150046 Sequence 150046, A
11	125	6.0	601	4	US-09-949-016-150029 Sequence 150029, A
12	121	5.8	601	4	US-09-949-016-150041 Sequence 150041, A
13	121	5.8	601	4	US-09-949-016-150042 Sequence 150042, A
14	121	5.8	601	4	US-09-949-016-150008 Sequence 150008, A
15	119	5.7	601	4	US-09-949-016-150055 Sequence 150055, A
16	119	5.7	601	4	US-09-949-016-150055 Sequence 150055, A
17	110	5.3	601	4	US-09-949-016-150032 Sequence 150032, A
18	94	4.5	244	4	US-09-471-276-495 Sequence 495, Appl
19	78	3.7	601	4	US-09-949-016-150007 Sequence 150007, A
20	76	3.6	601	4	US-09-949-016-150018 Sequence 150018, A
21	76	3.6	601	4	US-09-949-016-150018 Sequence 150018, A
22	76	3.6	601	4	US-09-949-016-150018 Sequence 150018, A
23	30	1.4	1681	4	US-09-023-655-453 Sequence 453, Appl
24	20	1.0	273	4	US-09-513-999C-14761 Sequence 14761, A
25	20	1.0	440	3	US-09-397-787-305 Sequence 305, Appl
26	20	1.0	444	4	US-09-621-976-14139 Sequence 14139, A
27	20	1.0	445	3	US-09-397-787-274 Sequence 274, Appl

c	28	20	1.0	174259	4	US-09-949-016-11968	Sequence 11968, A
c	29	20	1.0	174262	4	US-09-949-016-14259	Sequence 14259, A
c	30	19	0.9	169	1	US-08-166-346A-8	Sequence 8, Appl1
c	31	19	0.9	459	4	US-09-621-976-8324	Sequence 8324, Ap
c	32	19	0.9	3969	3	US-09-518-386B-4	Sequence 4, Appl1
c	33	19	0.9	4396	3	US-09-821-736-1	Sequence 1, Appl1
c	34	19	0.9	14721	4	US-09-949-016-13507	Sequence 13507, A
c	35	19	0.9	25199	4	US-09-949-016-13361	Sequence 13361, A
c	36	19	0.9	129658	4	US-09-949-016-17195	Sequence 17195, A
c	37	19	0.9	186734	4	US-09-949-016-14870	Sequence 14870, A
c	38	19	0.9	193689	4	US-09-949-016-12350	Sequence 12350, A
c	39	19	0.9	193689	4	US-09-949-016-13088	Sequence 13088, A
c	40	19	0.9	200663	4	US-09-949-016-12569	Sequence 12569, A
c	41	19	0.9	203093	4	US-09-949-016-14445	Sequence 14445, A
c	42	18	0.9	78	2	US-08-749-852-56	Sequence 56, Appl
c	43	18	0.9	78	2	US-08-749-852-58	Sequence 58, Appl
c	44	18	0.9	511	4	US-09-902-540-1374	Sequence 1374, Ap
c	45	18	0.9	531	4	US-09-252-991A-2223	Sequence 2223, Ap

#### ALIGNMENTS

RESULT 1									
US-09-318-448-23									
Sequence 23, Application US/09318448									
Patent No. 6210950									
GENERAL INFORMATION:									
APPLICANT: Johnson, William G.									
APPLICANT: Steenroos, Edward S.									
TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING									
FILE REFERENCE: 601-1-057									
CURRENT APPLICATION NUMBER: US/09/318,448									
CURRENT FILING DATE: 1999-05-25									
NUMBER OF SEQ ID NOS: 46									
SOFTWARE: Patentin Ver. 2.0									
SEQ ID NO 23									
LENGTH: 3259									
TYPE: DNA									
ORGANISM: Homo sapiens									
US-09-318-448-23									
Query Match									
Best Local Similarity 99.8%; Pred. No. 0;									
Matches 2093; Conservative 0; Mismatches 0; Indels 4; Gaps 1;									
QY	1	ATGAGGAGGTTCTGTTACTATATGCTACACAGCAGGAGCAAGGCAATCGCAGAA	60						
DB	80	ATGAGGAGGTTCTGTTACTATATGCTACACAGCAGGAGCAAGGCAATCGCAGAA	139						
QY	61	GAATGTGTGACAGAGCTGTGTATCTTGTGAGATCTTCACTGTATTAGTAA	120						
DB	140	GAATGTGTGACAGAGCTGTGTATCTTGTGAGATCTTCACTGTATTAGTAA	199						
QY	121	TCGGATTAATGACCTTAACAAACGCTCTCTGTGTGTGTGTTCTACACAG	180						
DB	200	TCGGATTAATGACCTTAACAAACGCTCTCTGTGTGTGTGTTCTACACAG	259						
QY	181	GGCAGCGAGAGCCACCCGACAGAGCCGAGTTGTTAAGAAATACAGAACAAAC	240						
DB	260	GGCAGCGAGAGCCACCCGACAGAGCCGAGTTGTTAAGAAATACAGAACAAAC	319						
QY	241	CTGCCGGTGAATTTCTTGTCTACCTCGGTATGGTTACTGGTCTCGGTATTCAGAA	300						
DB	320	CTGCCGGTGAATTTCTTGTCTACCTCGGTATGGTTACTGGTCTCGGTATTCAGAA	379						
QY	301	TACACTTACTTTTGCAATGGGGGGAAGATATTGATTAACGACTTGAGAGCC	360						
DB	380	TACACTTACTTTTGCAATGGGGGGAAGATATTGATTAACGACTTGAGAGCC	439						
QY	361	CGGATTTTATGACACTGACATGAGATGACTGTGTGTTAGAACTTGTTGAG	420						

Db 440 CCGCATTTCTATGACACTGACATGACAGATGACTGTGTAGGTTTGAACCTTTGGTTGAG 499  
Qy 421 CCGTGATGCTGGAATCTGCGCCAGCCCTCGAAGAGATTTAGTCAACAGAGACA 480  
Db 500 CCGTGATGCTGGAATCTGCGCCAGCCCTCGAAGAGATTTAGTCAACAGAGACA 559  
Qy 481 GAGGAGATTAAGTGGCGCACTCCGGTGGCATCACCTGCAATCTTTAGAGACAGACTTTG 540  
Db 560 GAGGAGATTAAGTGGCGCACTCCGGTGGCATCACCTGCAATCTTTAGAGACAGACTTTG 619  
Qy 541 AAGTCAGAGCTGCTACACATTTGAATCTCAAGTGGAGCTTGGAGATTCGATTTACAGA 600  
Db 620 AAGTCAGAGCTGCTACACATTTGAATCTCAAGTGGAGCTTGGAGATTCGATTTACAGA 679  
Qy 601 AGAAGGATTTCTGAGTTTGAAGCAAAATGCAAGTGAACCAACCAATCAATGTTGA 660  
Db 680 AGAAGGATTTCTGAGTTTGAAGCAAAATGCAAGTGAACCAACCAATCAATGTTGA 739  
Qy 661 ATTGAAGATTTGAGTCTCTCACTTACCCGTTCCGTTACCCCACTCTCAAGCCTCTG 720  
Db 740 ATTGAAGATTTGAGTCTCTCACTTACCCGTTCCGTTACCCCACTCTCAAGCCTCTG 799  
Qy 721 AATATTCCTGGTTTACCCCGAATATTTTACAGGTACATGTCGAGAGAGTCTCTTGGCAG 780  
Db 800 AATATTCCTGGTTTACCCCGAATATTTTACAGGTACATGTCGAGAGAGTCTCTTGGCAG 859  
Qy 781 GAGGAAAGCCAGATATCTGTGACTTCAGCAGATCCAGTTTTCAGAGTCCCAATTTCAAG 840  
Db 860 GAGGAAAGCCAGATATCTGTGACTTCAGCAGATCCAGTTTTCAGAGTCCCAATTTCAAG 919  
Qy 841 GCAGTTCACTTACTACGAATGATGCAATAAAACCACTGCTGTGTAATTTGACATT 900  
Db 920 GCAGTTCACTTACTACGAATGATGCAATAAAACCACTGCTGTGTAATTTGACATT 979  
Qy 901 TCAATTCAGACTTTTCTATGAGCCCTGAGATGCTTACGCGTGAATCTGCTTACAGT 960  
Db 980 TCAATTCAGACTTTTCTATGAGCCCTGAGATGCTTACGCGTGAATCTGCTTACAGT 1039  
Qy 961 GATTTGAGGTACAAAGCTTACTCCAAAGACTGCACTTGAAGTAAAGAGACACTGC 1020  
Db 1040 GATTTGAGGTACAAAGCTTACTCCAAAGACTGCACTTGAAGTAAAGAGACACTGC 1099  
Qy 1021 GTCTTTTGAATTAAGGACACACAAAGAGAGAGCTTACCCAGCATATA 1080  
Db 1100 GTCTTTTGAATTAAGGACACACAAAGAGAGAGCTTACCCAGCATATA 1159  
Qy 1081 CTGCGGGAGATGCTCTCCAGTTCATTTTACCTGGTGTGTAAGAACCGAGCAATTCCT 1140  
Db 1160 CTGCGGGAGATGCTCTCCAGTTCATTTTACCTGGTGTGTAAGAACCGAGCAATTCCT 1219  
Qy 1141 AAAAAGCATTTTTCGAGCCCTTGTGACTATACAGTACAGTGTGAAAAAGCGCAG 1200  
Db 1220 AAAAAGCATTTTTCGAGCCCTTGTGACTATACAGTACAGTGTGAAAAAGCGCAG 1279  
Qy 1201 CTACAGAGCTGTGCACTAAACAAAGGGGACGCGATTATAGCCGCTTGTACAGATGCC 1260  
Db 1280 CTACAGAGCTGTGCACTAAACAAAGGGGACGCGATTATAGCCGCTTGTACAGATGCC 1339  
Qy 1261 TGGGCGCTGTTGATCTCTCCCTGCTTCCCTTCCCTTCCGCAAGCAACACTCAGCTC 1320  
Db 1340 TGGGCGCTGTTGATCTCTCCCTGCTTCCCTTCCCTTCCGCAAGCAACACTCAGCTC 1399  
Qy 1321 CTGCTCAGACATCTTCTTAAACTTCAACCCAGACCATATTTGTTGCAAGCTCAAGTTTA 1380  
Db 1400 CTGCTCAGACATCTTCTTAAACTTCAACCCAGACCATATTTGTTGCAAGCTCAAGTTTA 1459  
Qy 1381 TTTTACCCAGGAAAGCTTCATTTTGTCTTCAACATTTGTAATTTCTGTCTACTGCCACA 1440  
Db 1460 TTTTACCCAGGAAAGCTTCATTTTGTCTTCAACATTTGTAATTTCTGTCTACTGCCACA 1519  
Qy 1441 ACAGAGGTTCTGCGGAGGAGATATGACAGGCTGCGCTGCTGTTGTTGTTGCTTCAATT 1500  
Db 1520 ACAGAGGTTCTGCGGAGGAGATATGACAGGCTGCGCTGCTGTTGTTGTTGCTTCAATT 1579

Qy 1501 CTTACGCCAATATACATGATCCCATGAAAGACAGCGGAAAGCCCTGGCTCTTAAGATA 1560  
Db 1580 CTTACGCCAATATACATGATCCCATGAAAGACAGCGGAAAGCCCTGGCTCTTAAGATA 1639  
Qy 1561 TCCATCTCTCTCTGAAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATTC 1620  
Db 1640 TCCATCTCTCTCTGAAACAAATTTCTTCACTTACAGATGACCCCTCAATCCCATTC 1699  
Qy 1621 ATATAGTGGGTCCAGGAACCGGATAGCCCGTTATTTGGGTTCTTACAACT---AG 1676  
Db 1700 ATATAGTGGGTCCAGGAACCGGATAGCCCGTTATTTGGGTTCTTACAACTTAGAGAG 1759  
Qy 1677 AAATCCAAAGAACCAACCCAGATGGAATTTTGAAGCATGTGTTGTTTGGCTGC 1736  
Db 1760 AAATCCAAAGAACCAACCCAGATGGAATTTTGAAGCATGTGTTGTTTGGCTGC 1819  
Qy 1737 AGGCAATAGATAGGATTTATCTATTCAGAAAAAGACTCAGACATTTCTTAAAGCTGG 1796  
Db 1820 AGGCAATAGATAGGATTTATCTATTCAGAAAAAGACTCAGACATTTCTTAAAGCTGG 1879  
Qy 1797 ATCTTACTCATCTAAAGGTTTCTCTCAAGAGATGCTCTGTGAGAGAGAGAAAGCC 1856  
Db 1880 ATCTTACTCATCTAAAGGTTTCTCTCAAGAGATGCTCTGTGAGAGAGAGAAAGCC 1939  
Qy 1857 CCAGCAAGTATGTCAGAACCAATCCAGCTTCATGCGCAGCAGGTGGCGAAGATCTTC 1916  
Db 1940 CCAGCAAGTATGTCAGAACCAATCCAGCTTCATGCGCAGCAGGTGGCGAAGATCTTC 1999  
Qy 1917 CTCACAGAAAGGCGCATTTTATGTGTGTGAGATGCAAAAGATTTGGCCAAAGATGTA 1976  
Db 2000 CTCACAGAAAGGCGCATTTTATGTGTGTGAGATGCAAAAGATTTGGCCAAAGATGTA 2059  
Qy 1977 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGGTTGAGTTGAAAACTAGAAAGCATG 2036  
Db 2060 CATGATGCCCTTGTGCAATTAATTAAGCAAGAGGTTGAGTTGAAAACTAGAAAGCATG 2119  
Qy 2037 AAAACCTGGCCACTTTAAAGAGAAAAAGCTACCTTCAAGATATTTGGTCATTA 2093  
Db 2120 AAAACCTGGCCACTTTAAAGAGAAAAAGCTACCTTCAAGATATTTGGTCATTA 2176

RESULT 2  
US-09-949-016-4215  
; Sequence 4215, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FaSTSeq for Windows Version 4.0  
; SEQ ID NO 4215  
; LENGTH: 3242  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-4215

Query Match 78.4%; Score 1640; DB 4; Length 3242;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 2090; Conservative 0; Mismatches 3; Indels 4; Gaps 1;  
Qy 1 ATGAGAGGTTTCTGTTACTATATGCTTACACAGAGGACAGGCAAGGCCATCGCAGAA 60  
|||||

80 ATGAGAGGTTTCTGTATCTATATGCTACACAGACAGGACAGGCAAGCCATCCGAGAA 139  
61 GAAATGTGACCAAGCTGTGTATCATGTATTTCTGCGAGCTTCACTGTATTTAGTAA 120  
140 GAAATATGTGACCAAGCTGTGTATCATGTATTTCTGCGAGCTTCACTGTATTTAGTAA 199  
121 TCCGATTAAGTATGACCTTAATAAACCCGAAACAGCTCTCTGTGTGTGTGTCTTCAACG 180  
200 TCCGATTAAGTATGACCTTAATAAACCCGAAACAGCTCTCTGTGTGTGTGTCTTCAACG 259  
181 GGCACCGAGAGACCAACCCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACA 240  
260 GGCACCGAGAGACCAACCCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACA 319  
241 CTGCGCGGTGATTTCTTTGCTCACTGCGGTATGTGGTTATCTGGGTCTGGGTATTCAGA 300  
320 CTGCGCGGTGATTTCTTTGCTCACTGCGGTATGTGGTTATCTGGGTCTGGGTATTCAGA 379  
301 TACACCTACTTTTGGCAATGGGGGAAAGATTAATTGAATAACGACTTCAAGAGCTGGAGCC 360  
380 TACACCTACTTTTGGCAATGGGGGAAAGATTAATTGAATAACGACTTCAAGAGCTGGAGCC 439  
361 CGGCAATTTCTATGACACTGACATGACATGATGATGTGTAGATTGAACTTGTGTAG 420  
440 CGGCAATTTCTATGACACTGACATGACATGATGATGTGTAGATTGAACTTGTGTAG 499  
421 CCGTGAGTTGCTGACCTGTGCGCAAGCTTCAAGAAACATTTTATGTTCAAGCAGAGACA 480  
500 CCGTGAGTTGCTGACCTGTGCGCAAGCTTCAAGAAACATTTTATGTTCAAGCAGAGACA 559  
481 GAGGAGATTAAGTGGGCACTCCCGGTGGATCACTCGATCCCTGAGGACAGACCTGTG 540  
560 GAGGAGATTAAGTGGGCACTCCCGGTGGATCACTCGATCCCTGAGGACAGACCTGTG 619  
541 AAGTCAGAGCTGCTACATGATGATCTCAAGTCAAGCTTCTGAGATTCTGATTTAGGA 600  
620 AAGTCAGAGCTGCTACATGATGATCTCAAGTCAAGCTTCTGAGATTCTGATTTAGGA 679  
601 AAGGAGATTTCTGAGGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATTCATTTGTA 660  
680 AAGGAGATTTCTGAGGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATTCATTTGTA 739  
661 AATGAGACTTTGAGGCTCACTTACCGGTGGATCCCGCACTCTCAAGGCTGTCTG 720  
740 AATGAGACTTTGAGGCTCACTTACCGGTGGATCCCGCACTCTCAAGGCTGTCTG 799  
721 AATATCTGCTGTTTACCCCAAGATATTTTACAGTATCTGCAAGAGTCTCTGTGCGAG 780  
800 AATATCTGCTGTTTACCCCAAGATATTTTACAGTATCTGCAAGAGTCTCTGTGCGAG 859  
781 GAGGAAAGCCAAAGTATCTGTGACTTCAAGAGATCCAGTTTCAAGTCCCAATTTCAAG 840  
860 GAGGAAAGCCAAAGTATCTGTGACTTCAAGAGATCCAGTTTCAAGTCCCAATTTCAAG 919  
841 GCAAGTCACTTACAGATGATGCTCAATTAACCACTCTGTGTGTATTTGAGCAAT 900  
920 GCAAGTCACTTACAGATGATGCTCAATTAACCACTCTGTGTGTATTTGAGCAAT 979  
901 TCAATATACAGACTTTTCTATGAGCTGAGATGCTTCAAGCTGTGATCTGCTTACAGT 960  
980 TCAATATACAGACTTTTCTATGAGCTGAGATGCTTCAAGCTGTGATCTGCTTACAGT 1039  
961 GATTTGAGGTACAAAGCTTCAAGAGCTGCAAGTGAATTAAGAGAGAGCACTGC 1020  
1040 GATTTGAGGTACAAAGCTTCAAGAGCTGCAAGTGAATTAAGAGAGAGCACTGC 1099  
1021 GTCTTTTGAATTAAGAGAGCAAAAGAAAGAGAGCTTCAAGTCCCAAGCAATTA 1080  
1100 GTCTTTTGAATTAAGAGAGCAAAAGAAAGAGAGCTTCAAGTCCCAAGCAATTA 1159  
1081 CTTGCGGAGATGTTCTCTCAAGTCAATTTTATCTGTGTCTTGAATTCAGAGCAATTTCT 1140  
1160 CTTGCGGAGATGTTCTCTCAAGTCAATTTTATCTGTGTCTTGAATTCAGAGCAATTTCT 1219

1141 AAAAGGCAATTTTTCGAGCCCTTGTGACTTAAACAGTGAACAGTGTGAAAAGCGCAG 1200  
1220 AAAAGGCAATTTTTCGAGCCCTTGTGACTTAAACAGTGAACAGTGTGAAAAGCGCAG 1279  
1201 CTACAGAGCTGTGCAATTAACAGAGGCAAGCCGATTAATGACCTTGTGTACAGATGCC 1260  
1280 CTACAGAGCTGTGCAATTAACAGAGGCAAGCCGATTAATGACCTTGTGTACAGATGCC 1339  
1261 TGTGCTGCTGTTGGAATCTCTGCTTCTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTG 1320  
1340 TGTGCTGCTGTTGGAATCTCTGCTTCTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTG 1399  
1321 CTGCTGCAACCTTCTTAACTTCAACCCAGACCAATTAATGCTGTGCAAGCTCAAGTTTA 1380  
1400 CTGCTGCAACCTTCTTAACTTCAACCCAGACCAATTAATGCTGTGCAAGCTCAAGTTTA 1459  
1381 TTTCAACCAAGAAAGCTCAATTTGTCTTCAACATTTGTGAAATTTCTGTCTACCTGCACA 1440  
1460 TTTCAACCAAGAAAGCTCAATTTGTCTTCAACATTTGTGAAATTTCTGTCTACCTGCACA 1519  
1441 ACAGAGTTCTGCGGAGAGGATATGTACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
1520 ACAGAGTTCTGCGGAGAGGATATGTACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1579  
1501 CTTGAGCCAAACATACATGATCCCATGAAAGCAGCGGAAAGCCCTGCTCTTAAGATA 1560  
1580 CTTGAGCCAAACATACATGATCCCATGAAAGCAGCGGAAAGCCCTGCTCTTAAGATA 1639  
1561 TCCATCTCTCTGCAACCAAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
1640 TCCATCTCTCTGCAACCAAAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1699  
1621 ATATATGTGGTTCAGAAACCGGCAATGCTGCTTATTTGGTCTCTTCAAT---AG 1676  
1700 ATATATGTGGTTCAGAAACCGGCAATGCTGCTTATTTGGTCTCTTCAATATGAGAG 1759  
1677 AAATCTCAAGAAACAACCCAGATGAAATTTTGGAGCAATGCTGCTTATTTGGCTGC 1736  
1760 AAATCTCAAGAAACAACCCAGATGAAATTTTGGAGCAATGCTGCTTATTTGGCTGC 1819  
1737 AGGCAATTAAGATGAGATTAATCTATTCAGAAAGAGCTCAGACATTTCTTAAAGCATG 1796  
1820 AGGCAATTAAGATGAGATTAATCTATTCAGAAAGAGCTCAGACATTTCTTAAAGCATG 1879  
1797 ATCTTAATCTCAATTAAGATTTCTTCTCAAGATGCTCTGTTGGGAGAGAGAGCC 1856  
1880 ATCTTAATCTCAATTAAGATTTCTTCTCAAGATGCTCTGTTGGGAGAGAGAGCC 1939  
1857 CCAGCAAGATGATTAAGAAACAACCCAGATGAAATTTTGGAGCAATGCTCTC 1916  
1940 CCAGCAAGATGATTAAGAAACAACCCAGATGAAATTTTGGAGCAATGCTCTC 1999  
1917 CTCAGAGAGAGCGCCATTTATGTGTGTGAGATGCAAGAAATATGCTCAAGATGTA 1976  
2000 CTCAGAGAGAGCGCCATTTATGTGTGTGAGATGCAAGAAATATGCTCAAGATGTA 2059  
1977 CATGATGCTCTTGTGCAATTAATTAAGCAAAAGGTTGAGTTGAAAACCTAAGCAATG 2036  
2060 CATGATGCTCTTGTGCAATTAATTAAGCAAAAGGTTGAGTTGAAAACCTAAGCAATG 2119  
2037 AAAACCTGCGCACTTTAAAGAAAGAAAGAAAGCTTCAAGATTTTGGTCAATTA 2093  
2120 AAAACCTGCGCACTTTAAAGAAAGAAAGAAAGCTTCAAGATTTTGGTCAATTA 2176

RESULT 3  
US-08-905-223-71  
Sequence 71, Application US/08905223  
Patent No. 6222029  
GENERAL INFORMATION:  
APPLICANT: Edwards, Jean-Baptiste D.  
APPLICANT: Duclercq, Aymeric

APPLICANT: Lacroix, Bruno  
 TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS  
 NUMBER OF SEQUENCES: 503  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Knobbe, Martens, Olson & Bear  
 STREET: 501 West Broadway  
 CITY: San Diego  
 STATE: California  
 COUNTRY: USA  
 ZIP: 92101-3505  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: Win95  
 SOFTWARE: Word  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/905,223  
 FILING DATE:  
 CLASSIFICATION: 536  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Israel, Ned A.  
 REGISTRATION NUMBER: 29,655  
 REFERENCE/DOCKET NUMBER:  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (619) 235-8550  
 TELEFAX: (619) 235-0176  
 INFORMATION FOR SEQ ID NO: 71:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 390 base pairs  
 TYPE: NUCLEIC ACID  
 STRANDEDNESS: DOUBLE  
 TOPOLOGY: LINEAR  
 MOLECULE TYPE: CDNA  
 ORIGINAL SOURCE:  
 ORGANISM: Homo Sapiens  
 TISSUE TYPE: Brain  
 FEATURE:  
 NAME/KEY: s1g\_peptide  
 LOCATION: 289..357  
 IDENTIFICATION METHOD: Von Heijne matrix  
 OTHER INFORMATION: score 6.9  
 OTHER INFORMATION: seq SLSLASHSVSC/SN  
 US-08-905-223-71

Query Match 18.4%; Score 386; DB 3; Length 390;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-188;  
 Matches 386; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 970 GTACAAAGCTTCTCCAAAGACTGCGAGCTTGAAGATAAAAGAGCAGCTGCGTCTTTTG 1029  
 DB 3 GTACAAAGCTTCTCCAAAGACTGCGAGCTTGAAGATAAAAGAGCAGCTGCGTCTTTTG 62  
 QY 1030 AAATAAAGGAGAGCAAGAAAGAAAGAGCTACCTTACCAGCATATACCGGAGGGA 1089  
 DB 63 AAATAAAGGAGAGCAAGAAAGAAAGAGCTACCTTACCAGCATATACCGGAGGGA 122  
 QY 1090 TGTTCCTCCAGTTCATTTTACCTGTGTCTTGAATTCGAGCAATTCCTAAAAAGGCA 1149  
 DB 123 TGTTCCTCCAGTTCATTTTACCTGTGTCTTGAATTCGAGCAATTCCTAAAAAGGCA 182  
 QY 1150 TTTTTCGAGAGCCCTTGTGAGCTATACAGTGAAGCTGCTGAAAAAGCCGAGGCTACAGAG 1209  
 DB 183 TTTTTCGAGAGCCCTTGTGAGCTATACAGTGAAGCTGCTGAAAAAGCCGAGGCTACAGAG 242  
 QY 1210 CTGTCAGTAAACAAAGGGGAGCCGATTAAGCCGCTTTTATAGAGATGCTGTGCTGCTGC 1269  
 DB 243 CTGTCAGTAAACAAAGGGGAGCCGATTAAGCCGCTTTTATAGAGATGCTGTGCTGCTGC 302  
 QY 1270 TTGTTGATCTGAA 1329  
 DB 303 TTGTTGATCTGAA 362  
 QY 1330 CATCTTCTTAACCTTCAACCCAGACC 1355

Db 363 CATCTTCTTAACCTTCAACCCAGACC 388  
 RESULT 4  
 US-09-949-016-150019  
 Sequence 150019, Application US/09949016  
 Patent No. 6812339  
 GENERAL INFORMATION:  
 APPLICANT: VENTER, J. Craig et al.  
 TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 FILE REFERENCE: CL001307  
 CURRENT APPLICATION NUMBER: US/09/949,016  
 PRIOR FILING DATE: 2000-04-14  
 PRIOR APPLICATION NUMBER: 60/241,755  
 PRIOR FILING DATE: 2000-10-20  
 PRIOR APPLICATION NUMBER: 60/237,768  
 PRIOR FILING DATE: 2000-10-03  
 PRIOR APPLICATION NUMBER: 60/231,498  
 PRIOR FILING DATE: 2000-09-08  
 NUMBER OF SEQ ID NOS: 207012  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO: 150019  
 LENGTH: 601  
 TYPE: DNA  
 ORGANISM: Human  
 US-09-949-016-150019

Query Match 15.8%; Score 330; DB 4; Length 601;  
 Best Local Similarity 99.7%; Pred. No. 1.8e-159;  
 Matches 380; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 401 GTTTAGAACTTGTGTTGAGCCGTGATGCTGACTCTGCGCAGCCCTCAGAAAGCAAT 460  
 DB 178 GTTTAGAACTTGTGTTGAGCCGTGATGCTGACTCTGCGCAGCCCTCAGAAAGCAAT 237  
 QY 461 TTAGTCAAGCAGAGAGCAAGAGAGATAGTGGCGCACTCCCGTGGCATCCTTGAT 520  
 DB 238 TTAGTCAAGCAGAGAGCAAGAGAGATAGTGGCGCACTCCCGTGGCATCCTTGAT 297  
 QY 521 CCTTGAAGCAGACCTTGTGAAGTCAAGCTGCTCACTTAATTAATTCAGTGCAGCTTC 580  
 DB 298 CCTTGAAGCAGACCTTGTGAAGTCAAGCTGCTCACTTAATTAATTCAGTGCAGCTTC 357  
 QY 581 TGAGATTCATGATTCAGAGAAAGAAAGATTCGAGGTTTGAAGCAAAATGCAAGTAA 640  
 DB 358 TGAGATTCATGATTCAGAGAAAGAAAGATTCGAGGTTTGAAGCAAAATGCAAGTAA 417  
 QY 641 GCAACCAATCCAATGTTGTAATTGAAGTGAAGTCTCACTTACCCGTTGGTACCCC 700  
 DB 418 GCAACCAATCCAATGTTGTAATTGAAGTGAAGTCTCACTTACCCGTTGGTACCCC 477  
 QY 701 CACTTCACAGAGCTCTGGAATATTCCTGTTTACCCCAAGAAATTTACAGTGCATC 760  
 DB 478 CACTTCACAGAGCTCTGGAATATTCCTGTTTACCCCAAGAAATTTACAGTGCATC 537  
 QY 761 TGCAGAGTCTCTTGGCAGG 781  
 DB 538 TGCAGAGTCTCTTGGCAGG 558  
 RESULT 5  
 US-09-949-016-15957  
 Sequence 15957, Application US/09949016  
 Patent No. 6812339  
 GENERAL INFORMATION:  
 APPLICANT: VENTER, J. Craig et al.  
 TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 FILE REFERENCE: CL001307  
 CURRENT APPLICATION NUMBER: US/09/949,016  
 CURRENT FILING DATE: 2000-04-14



PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 15957  
LENGTH: 35916  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-15957

Query Match 15.8%; Score 330; DB 4; Length 35916;  
Best Local Similarity 99.7%; Pred. No. 2.3e-159;  
Matches 380; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 401 GTTTGAACTTGTGGTGGCCGTGATTTGCTGACTCTGGCCAGCCCTCAGAAAGCATT 460  
DB 10781 GTTTGAACTTGTGGTGGCCGTGATTTGCTGACTCTGGCCAGCCCTCAGAAAGCATT 10840  
QY 461 TTAGTCAAGCAGAGACAAGAGATTAAGTGGCCACTCCCGGTGGCATCACTGCAT 520  
DB 10841 TTAGTCAAGCAGAGACAAGAGATTAAGTGGCCACTCCCGGTGGCATCACTGCAT 10900  
QY 521 CCTTGAGACAGACCTTGTGAAGTCAAGCTGCTACATTTGAATTTCAAGTCAGCTTC 580  
DB 10901 CCTTGAGACAGACCTTGTGAAGTCAAGCTGCTACATTTGAATTTCAAGTCAGCTTC 10960  
QY 581 TGAGATTCATGATTCAGAGAAAGAAAGATTCTGAGTTTGAAGCAAAATGACAGTACA 640  
DB 10961 TGAGATTCATGATTCAGAGAAAGAAAGATTCTGAGTTTGAAGCAAAATGACAGTACA 11020  
QY 641 GCAACCAATCCAAATGTTGAATTAAGTGAAGTCTCACTTACCCGTTGGTACCCC 700  
DB 11021 GCAACCAATCCAAATGTTGAATTAAGTGAAGTCTCACTTACCCGTTGGTACCCC 11080  
QY 701 CACTCTCAAGCCTCTCTGATATTTCTGTTTACCCCGAATATTTACAGATACATC 760  
DB 11081 CACTCTCAAGCCTCTCTGATATTTCTGTTTACCCCGAATATTTACAGATACATC 11140  
QY 761 TGCAGAGTCTCTGGCCAGG 781  
DB 11141 TGCAGAGTCTCTGGCCAGG 11161

## RESULT 6

US-09-949-016-150020  
Sequence 150020, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 150020  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-150020

Query Match 13.3%; Score 279; DB 4; Length 601;

Best Local Similarity 99.5%; Pred. No. 3.5e-133;  
Matches 379; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 401 GTTTGAACTTGTGGTGGCCGTGATTTGCTGACTCTGGCCAGCCCTCAGAAAGCATT 460  
DB 165 GTTTGAACTTGTGGTGGCCGTGATTTGCTGACTCTGGCCAGCCCTCAGAAAGCATT 224  
QY 461 TTAGTCAAGCAGAGACAAGAGATTAAGTGGCCACTCCCGGTGGCATCACTGCAT 520  
DB 225 TTAGTCAAGCAGAGACAAGAGATTAAGTGGCCACTCCCGGTGGCATCACTGCAT 284  
QY 521 CCTTGAGACAGACCTTGTGAAGTCAAGCTGCTACATTTGAATTTCAAGTCAGCTTC 580  
DB 285 CCTTGAGACAGACCTTGTGAAGTCAAGCTGCTACATTTGAATTTCAAGTCAGCTTC 344  
QY 581 TGAGATTCATGATTCAGAGAAAGAAAGATTCTGAGTTTGAAGCAAAATGACAGTACA 640  
DB 345 TGAGATTCATGATTCAGAGAAAGAAAGATTCTGAGTTTGAAGCAAAATGACAGTACA 404  
QY 641 GCAACCAATCCAAATGTTGAATTAAGTGAAGTCTCACTTACCCGTTGGTACCCC 700  
DB 405 GCAACCAATCCAAATGTTGAATTAAGTGAAGTCTCACTTACCCGTTGGTACCCC 464  
QY 701 CACTCTCAAGCCTCTCTGATATTTCTGTTTACCCCGAATATTTACAGATACATC 760  
DB 465 CACTCTCAAGCCTCTCTGATATTTCTGTTTACCCCGAATATTTACAGATACATC 524  
QY 761 TGCAGAGTCTCTGGCCAGG 781  
DB 525 TGCAGAGTCTCTGGCCAGG 545

## RESULT 7

US-09-949-016-150037  
Sequence 150037, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 150037  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-150037

Query Match 9.0%; Score 189; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 8.6e-87;  
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1369 AGCTCAAGTTATTTTCAACCCGGAAGCTCCATTTTGTCTTCAACATTTGTGAATTTCTG 1428  
DB 18 AGCTCAAGTTATTTTCAACCCGGAAGCTCCATTTTGTCTTCAACATTTGTGAATTTCTG 77  
QY 1429 TCTACTGCCACAAGAGGTTCTGGGAAAGGAGATATGACAGGCTGGCTGCTGTGTTG 1488  
DB 78 TCTACTGCCACAAGAGGTTCTGGGAAAGGAGATATGACAGGCTGGCTGCTGTGTTG 137  
QY 1489 GTTGCTTCAAGTTCTTCAAGCAATACATGATCCCATGAAAGACGGGGAAGCCCTTG 1548  
DB 138 GTTGCTTCAAGTTCTTCAAGCAATACATGATCCCATGAAAGACGGGGAAGCCCTTG 197

QY 1549 GCTCCTTAG 1557  
Db 198 GCTCCTTAG 206

## RESULT 8

US-09-566-921-88  
Sequence 89, Application US/09566921  
Patent No. 6682888  
GENERAL INFORMATION:  
APPLICANT: Loring, Jeanne F.  
APPLICANT: Tingley, Debora W.  
APPLICANT: Edwards, Carla M.  
TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE  
FILE REFERENCE: PA-0024 US  
CURRENT APPLICATION NUMBER: US/09/566,921  
CURRENT FILING DATE: 2000-05-05  
NUMBER OF SEQ ID NOS: 138  
SOFTWARE: PERL Program  
SEQ ID NO 88  
LENGTH: 2475  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. 6682888 255828.26  
LOCATION: 1001, 1011  
OTHER INFORMATION: a, t, c, g, or other  
US-09-566-921-88

Query Match 7.5%; Score 158; DB 4; Length 2475;  
Best Local Similarity 100.0%; Pred. No. 8.9e-71;  
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 525 GAGACAGACCTTGTGAGTCTGACAGCTGTACACATTAATCTCAAGTCAGCTTCTGAG 584  
Db 16 GAGACAGACCTTGTGAGTCTGACAGCTGTACACATTAATCTCAAGTCAGCTTCTGAG 75  
QY 585 ATTGATGATTCAGAGAGAGAGATTCAGGTTTGAAGCAAAATGCACTGAGCAAGCA 644  
Db 76 ATTGATGATTCAGAGAGAGAGATTCAGGTTTGAAGCAAAATGCACTGAGCAAGCA 135  
QY 645 CCAATCCAAATGTTGATTAATGAAGCTTTGAGTCTGAC 682  
Db 136 CCAATCCAAATGTTGATTAATGAAGCTTTGAGTCTGAC 173

## RESULT 9

US-09-949-016-150030  
Sequence 150030, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 150030  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-150030

Query Match 7.4%; Score 155; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 2.9e-69;  
Matches 155; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 904 AATACAGACTTTTCCATCAGCTGAGAGATGCTTCAGCGTATCTGCCCTTAACAGTGAT 963  
Db 320 AATACAGACTTTTCCATCAGCTGAGAGATGCTTCAGCGTATCTGCCCTTAACAGTGAT 379  
QY 964 TCTGAGTCAAAAGCTTACTCCAAAGACTGACGCTTGAAGATTAAGAGAGCACTGCGTC 1023  
Db 380 TCTGAGTCAAAAGCTTACTCCAAAGACTGACGCTTGAAGATTAAGAGAGCACTGCGTC 439  
QY 1024 CTTTGAATAATAAGGACAGACAAAGAGAAAGG 1058  
Db 440 CTTTGAATAATAAGGACAGACAAAGAGAAAGG 474

## RESULT 10

US-09-949-016-150031  
Sequence 150031, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 150031  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-150031

Query Match 6.9%; Score 145; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 4.1e-64;  
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 904 AATACAGACTTTTCCATCAGCTGAGAGATGCTTCAGCGTATCTGCCCTTAACAGTGAT 963  
Db 156 AATACAGACTTTTCCATCAGCTGAGAGATGCTTCAGCGTATCTGCCCTTAACAGTGAT 215  
QY 964 TCTGAGTCAAAAGCTTACTCCAAAGACTGACGCTTGAAGATTAAGAGAGCACTGCGTC 1023  
Db 216 TCTGAGTCAAAAGCTTACTCCAAAGACTGACGCTTGAAGATTAAGAGAGCACTGCGTC 275  
QY 1024 CTTTGAATAATAAGGACAGACAA 1048  
Db 276 CTTTGAATAATAAGGACAGACAA 300

## RESULT 11

US-09-949-016-150046  
Sequence 150046, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768

;; PRIOR FILING DATE: 2000-10-03  
;; PRIOR APPLICATION NUMBER: 60/231,498  
;; PRIOR FILING DATE: 2000-09-08  
;; NUMBER OF SEQ ID NOS: 207012  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 150046  
;; LENGTH: 601  
;; TYPE: DNA  
;; ORGANISM: Human  
US-09-949-016-150046

Query Match 6.5%; Score 137; DB 4; Length 601;  
Best Local Similarity 99.5%; Pred. No. 5.5e-60;  
Matches 187; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1761 TTCAGAAAAGACTGAGATTTCTTAAGCATGGAGTCTTAATCTTAAGTTTC 1820  
Db 413 TTCAGAAAAGACTGAGATTTCTTAAGCATGGAGTCTTAATCTTAAGTTTC 472  
Qy 1821 TTCTCAAGATGCTCTGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1880  
Db 473 TTCTCAAGATGCTCTGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 532  
Qy 1881 ATCCAGCTTCATGGCAGAGGTGGGAGAGATCTTCTCCAGAGAGAGCGCATTTAT 1940  
Db 533 ATCCAGCTTCATGGCAGAGGTGGGAGAGATCTTCTCCAGAGAGAGCGCATTTAT 592  
Qy 1941 GTGTGTGG 1948  
Db 593 GTGTGTGG 600

## RESULT 12

US-09-949-016-150047  
; Sequence 150047, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 150047  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-150047

Query Match 6.5%; Score 137; DB 4; Length 601;  
Best Local Similarity 99.5%; Pred. No. 5.5e-60;  
Matches 187; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1761 TTCAGAAAAGACTGAGATTTCTTAAGCATGGAGTCTTAATCTTAAGTTTC 1820  
Db 191 TTCAGAAAAGACTGAGATTTCTTAAGCATGGAGTCTTAATCTTAAGTTTC 250  
Qy 1821 TTCTCAAGATGCTCTGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1880  
Db 251 TTCTCAAGATGCTCTGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 310  
Qy 1881 ATCCAGCTTCATGGCAGAGGTGGGAGAGATCTTCTCCAGAGAGAGCGCATTTAT 1940  
Db 311 ATCCAGCTTCATGGCAGAGGTGGGAGAGATCTTCTCCAGAGAGAGCGCATTTAT 370

Qy 1941 GTGTGTGG 1948  
Db 371 GTGTGTGG 378

## RESULT 13

US-09-949-016-150029  
; Sequence 150029, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 150029  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-150029

Query Match 6.0%; Score 125; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 8.4e-54;  
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 779 AGAGGAAAGCCAGATATCTGTGACTTCAGAGATCCAGTTTCAAGTCCCAATTCA 838  
Db 379 AGAGGAAAGCCAGATATCTGTGACTTCAGAGATCCAGTTTCAAGTCCCAATTCA 438  
Qy 839 AGGAGTTCACTTCTAGCATGATGCTTAATAAACCACTGTGTGAATTTGACA 898  
Db 439 AGGAGTTCACTTCTAGCATGATGCTTAATAAACCACTGTGTGAATTTGACA 498  
Qy 899 TTTCA 903  
Db 499 TTTCA 503

## RESULT 14

US-09-949-016-150041  
; Sequence 150041, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 150041  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-150041

Query Match 5.8%; Score 121; DB 4; Length 601;

Best Local Similarity 100.0%; Pred. No. 9.7e-52;  
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1556 AGATATCCATCTCTCTCGAACAACAATCTTTCACCTTACGATGACCCCTCAATCC 1615  
|  
DB 124 AGATATCCATCTCTCTCGAACAACAATCTTTCACCTTACGATGACCCCTCAATCC 183

QY 1616 CCATCATTAATGGTGGGTCCAGAAACCGGCATAGCCCGTTTATTGGGTCTTACAACATA 1675  
|  
DB 184 CCATCATTAATGGTGGGTCCAGAAACCGGCATAGCCCGTTTATTGGGTCTTACAACATA 243

QY 1676 G 1676  
|  
DB 244 G 244

## RESULT 15

US-09-949-016-150042  
; Sequence 150042, Application US/09949016  
; Patent No. 681239  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CU001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 150042  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-150042

Query Match 5.8%; Score 121; DB 4; Length 601;

Best Local Similarity 100.0%; Pred. No. 9.7e-52;  
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1556 AGATATCCATCTCTCTCGAACAACAATCTTTCACCTTACGATGACCCCTCAATCC 1615  
|  
DB 95 AGATATCCATCTCTCTCGAACAACAATCTTTCACCTTACGATGACCCCTCAATCC 154

QY 1616 CCATCATTAATGGTGGGTCCAGAAACCGGCATAGCCCGTTTATTGGGTCTTACAACATA 1675  
|  
DB 155 CCATCATTAATGGTGGGTCCAGAAACCGGCATAGCCCGTTTATTGGGTCTTACAACATA 214

QY 1676 G 1676  
|  
DB 215 G 215

Search completed: August 27, 2005, 16:18:22  
Job time : 237.308 secs



Db 1 ATGAGAGGTTTCTGTTACTATATGCTACACAGAGGACAGGCAAAAGCCATCGCAGAA 60  
 QY 61 GAAATGTGTAGAGAGCTGTGTACATGATTTTCTGCAGATCTTCACTGTATTAGTAA 120  
 Db 61 GAAATGTGTAGAGAGCTGTGTACATGATTTTCTGCAGATCTTCACTGTATTAGTAA 120  
 QY 121 TCCGATTAAGTATGACCTTAAACCCGAAACAGCTCTCTTGTGTGTGTGTCTTACAG 180  
 Db 121 TCCGATTAAGTATGACCTTAAACCCGAAACAGCTCTCTTGTGTGTGTGTCTTACAG 180  
 QY 181 GGCACCGAGACCAACCCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCAAC 240  
 Db 181 GGCACCGAGACCAACCCGACACAGCCCGCAAGTTTGTAAAGAAATACAGAACCAAC 240  
 QY 241 CTGCGGTTGATTTCTTGTCTCACTGCGGTATGAGTTTCTGCGGTCTCGGTATTCAGAA 300  
 Db 241 CTGCGGTTGATTTCTTGTCTCACTGCGGTATGAGTTTCTGCGGTCTCGGTATTCAGAA 300  
 QY 301 TACACCTACTTTTGCATTTGGGGGGAAGATTAATGATTAACGACTTCAAGAGCTTGAAGCC 360  
 Db 301 TACACCTACTTTTGCATTTGGGGGGAAGATTAATGATTAACGACTTCAAGAGCTTGAAGCC 360  
 QY 361 CCGCATTTCTATGACATCTGACATGAGATGACTGTGTAGTTTGAACCTTGTGTGAG 420  
 Db 361 CCGCATTTCTATGACATCTGACATGAGATGACTGTGTAGTTTGAACCTTGTGTGAG 420  
 QY 421 CCGTGAATGCTGACATCTGCGCACCCCTCAGAAAGCAATTTTAAAGTCAACAGAGACA 480  
 Db 421 CCGTGAATGCTGACATCTGCGCACCCCTCAGAAAGCAATTTTAAAGTCAACAGAGACA 480  
 QY 481 GAGAGATTAAGTGGCGGACCTCCGCTGACATCTGCAATCTTGAAGACAGACCTTGTG 540  
 Db 481 GAGAGATTAAGTGGCGGACCTCCGCTGACATCTGCAATCTTGAAGACAGACCTTGTG 540  
 QY 541 AAGTCAGAGCTGCTACATCTGAAATCTCAAGTCAAGCTTCTGAGATTCATGATTCAGAA 600  
 Db 541 AAGTCAGAGCTGCTACATCTGAAATCTCAAGTCAAGCTTCTGAGATTCATGATTCAGAA 600  
 QY 601 AGAAGAGATTTCTGAGGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATCAATGTTGTA 660  
 Db 601 AGAAGAGATTTCTGAGGTTTGAAGCAAAATGCAAGTGAACAGCAACCAATCAATGTTGTA 660  
 QY 661 ATTGAAGATTTTGAAGTCTCACTTACCCGTTCCGTACCCCACTCTCAAGAGCTCTCTG 720  
 Db 661 ATTGAAGATTTTGAAGTCTCACTTACCCGTTCCGTACCCCACTCTCAAGAGCTCTCTG 720  
 QY 721 AATATCTGATTTACCCCAAGAAATTTTACAGGTAATCTGACAGAGTCTCTTGGCCAG 780  
 Db 721 AATATCTGATTTACCCCAAGAAATTTTACAGGTAATCTGACAGAGTCTCTTGGCCAG 780  
 QY 781 GAGGAAAGCAAGTATCTGTGACTTTCAGAGATTCAGATTTTTCAGTGCATTTTCAAG 840  
 Db 781 GAGGAAAGCAAGTATCTGTGACTTTCAGAGATTCAGATTTTTCAGTGCATTTTCAAG 840  
 QY 841 GCGATTTCACTTACTAGAAATGATGCAATTAACCACTCTGCTGGTGAATTTGACAT 900  
 Db 841 GCGATTTCACTTACTAGAAATGATGCAATTAACCACTCTGCTGGTGAATTTGACAT 900  
 QY 901 TCAAAATACAGACTTTTCTATCAGCTGTGAGATGCTTTCAGCGTGTGATCTGCTTACAG 960  
 Db 901 TCAAAATACAGACTTTTCTATCAGCTGTGAGATGCTTTCAGCGTGTGATCTGCTTACAG 960  
 QY 961 GATTTCTAGGTAACAAAGCTTCTCAAAAGACTGCAAGTTTGAAGTAAGAGAGCACTGC 1020  
 Db 961 GATTTCTAGGTAACAAAGCTTCTCAAAAGACTGCAAGTTTGAAGTAAGAGAGCACTGC 1020  
 QY 1021 GTCTTTTGAAGAAATTAAGGACAGACAAAGAAAGAGCTTACCTTACCCAGCATATA 1080  
 Db 1021 GTCTTTTGAAGAAATTAAGGACAGACAAAGAAAGAGCTTACCTTACCCAGCATATA 1080  
 QY 1081 CCGTGGAGATGTTCTCTCAAGTCAATTTTACCTGTGTCTTGAAGATCCAGCAATTTCT 1140  
 Db 1081 CCGTGGAGATGTTCTCTCAAGTCAATTTTACCTGTGTCTTGAAGATCCAGCAATTTCT 1140

QY 1141 AAAAGGCAATTTTGGAGCCCTTGTGACATTAACAGTGAACGTGTGAAAAAGCCAG 1200  
 Db 1141 AAAAGGCAATTTTGGAGCCCTTGTGACATTAACAGTGAACGTGTGAAAAAGCCAG 1200  
 QY 1201 CTACAGAGCTGTGACAGTAAACAGAGGAGCCGATTAATAGCCGCTTGTACAGATG 1260  
 Db 1201 CTACAGAGCTGTGACAGTAAACAGAGGAGCCGATTAATAGCCGCTTGTACAGATG 1260  
 QY 1261 TGTGCTGTGTGTGATCTCTCTCTGCTTCCCTTCTTGCAGCACTCACTAGTCTC 1320  
 Db 1261 TGTGCTGTGTGTGATCTCTCTCTGCTTCCCTTCTTGCAGCACTCACTAGTCTC 1320  
 QY 1321 CTGCTGAAATCTTCTTAATCTTCAACCCAGACATATTGTGTGCAAGCTCAAGTTA 1380  
 Db 1321 CTGCTGAAATCTTCTTAATCTTCAACCCAGACATATTGTGTGCAAGCTCAAGTTA 1380  
 QY 1381 TTTCAACCGAGAAAGCTCAATTTTGTCTTCAATTTGTGGAATTTCTGTCTAC 1440  
 Db 1381 TTTCAACCGAGAAAGCTCAATTTTGTCTTCAATTTGTGGAATTTCTGTCTAC 1440  
 QY 1441 ACAGAGTTCTGCGAGAGGAGTATGTACAGGCTGCGCTGTGTGTGCTTCAAGTT 1500  
 Db 1441 ACAGAGTTCTGCGAGAGGAGTATGTACAGGCTGCGCTGTGTGTGCTTCAAGTT 1500  
 QY 1501 CTTCAAGCTCAATCATGATCTCCATGAAACAGCGGAAAGCCCTGCTCTTAAGATA 1560  
 Db 1501 CTTCAAGCTCAATCATGATCTCCATGAAACAGCGGAAAGCCCTGCTCTCTTAAGATA 1560  
 QY 1561 TCCATCTCTCTGAAACCAATTTCTTCACTTACAGAGATGACCCCTCAATCCCATC 1620  
 Db 1561 TCCATCTCTCTGAAACCAATTTCTTCACTTACAGAGATGACCCCTCAATCCCATC 1620  
 QY 1621 ATATAGTGTGATTCAGAAACCGGACATAGCCCGTTATTTGAGTCTTACAAATAGAAAC 1680  
 Db 1621 ATATAGTGTGATTCAGAAACCGGACATAGCCCGTTATTTGAGTCTTACAAATAGAAAC 1680  
 QY 1681 TCCAGAACCAACACCCAGATGGAATTTTGGAGCAATGTGTGTTTGTGTGCTGAGGC 1740  
 Db 1681 TCCAGAACCAACACCCAGATGGAATTTTGGAGCAATGTGTGTTTGTGTGCTGAGGC 1740  
 QY 1741 ATAGGATAGGAGTTATCTAATTCAGAAAGAGTCAATTTCTTAAAGATGGAATCT 1800  
 Db 1741 ATAGGATAGGAGTTATCTAATTCAGAAAGAGTCAATTTCTTAAAGATGGAATCT 1800  
 QY 1801 TAACTCATTAAGGTTTCTCTTCAAGAGATGCTCTGTGTGAGGAGAGAAAGCCAG 1860  
 Db 1801 TAACTCATTAAGGTTTCTCTTCAAGAGATGCTCTGTGTGAGGAGAGAAAGCCAG 1860  
 QY 1861 CAAAGTATGACAGCAACATTCAGCTTATGCGCAGAGGCTGAGAAATCTCTCTC 1920  
 Db 1861 CAAAGTATGACAGCAACATTCAGCTTATGCGCAGAGGCTGAGAAATCTCTCTC 1920  
 QY 1921 AGGAGAACGGCAATTTATGTGTGTGAGATGCAAGAAATTTGCAAGATGTATCATG 1980  
 Db 1921 AGGAGAACGGCAATTTATGTGTGTGAGATGCAAGAAATTTGCAAGATGTATCATG 1980  
 QY 1981 ATGCTCTGTGCAATTAATTAAGCAAGAGTGTGAGTTGAAAACTTGAAGCATGAAAA 2040  
 Db 1981 ATGCTCTGTGCAATTAATTAAGCAAGAGTGTGAGTTGAAAACTTGAAGCATGAAAA 2040  
 QY 2041 CCTGGCACTTTAAAGAAAGAAAGGCTTACCTTCAAGATATTGTGTATTA 2093  
 Db 2041 CCTGGCACTTTAAAGAAAGAAAGGCTTACCTTCAAGATATTGTGTATTA 2093

RESULT 2  
 US-09-371-347-1  
 ; Sequence 1, Application US/09371347  
 ; Publication No. US20030082676A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Roy A. Gravel et al.  
 ; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE REDUCTASE.

TITLE OF INVENTION: CLONING, AND METHODS FOR EVALUATING RISK OF NEURAL TUBE  
FILE OF INVENTION: DEFECTS CARDIOVASCULAR DISEASE, AND CANCER  
FILE REFERENCE: 50004/003003  
CURRENT APPLICATION NUMBER: US/09/371,347  
CURRENT FILING DATE: 1999-08-10  
PRIOR APPLICATION NUMBER: 60/071,622  
PRIOR FILING DATE: 1998-01-16  
PRIOR APPLICATION NUMBER: 09/232,028  
PRIOR FILING DATE: 1999-01-15  
NUMBER OF SEQ ID NOS: 51  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 2097  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-371-347-1

Query Match 85.7%; Score 1793; DB 10; Length 2097;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 2093; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 1 ATGAGAGGTTCTGTTACTATATGCTACACAGGAGGAGGCAAGGCCATCGCAGAA 60  
DB 1 ATGAGAGGTTCTGTTACTATATGCTACACAGGAGGAGGCAAGGCCATCGCAGAA 60  
QY 61 GAAATGTGTGACCAAGCTGTGTGACATGATTTCTGAGATCTTCACTGATTTAGTGA 120  
DB 61 GAAATGTGTGACCAAGCTGTGTGACATGATTTCTGAGATCTTCACTGATTTAGTGA 120  
QY 121 TCCGATTAAGTATGACTTAATAAACCAGAAAGCTCTCTGTTGTTGTTGTTCTACAG 180  
DB 121 TCCGATTAAGTATGACTTAATAAACCAGAAAGCTCTCTGTTGTTGTTGTTCTACAG 180  
QY 181 GGCACCGGAGACCCACCCGACACAGCCCGCAAGTTGTTAAGAAATACAGAACCAACA 240  
DB 181 GGCACCGGAGACCCACCCGACACAGCCCGCAAGTTGTTAAGAAATACAGAACCAACA 240  
QY 241 CTGCGGTTGATTTCTTCTGCTACCTGCGGATGAGTTAAGTTCTGCTGATTCAGAA 300  
DB 241 CTGCGGTTGATTTCTTCTGCTACCTGCGGATGAGTTAAGTTCTGCTGATTCAGAA 300  
QY 301 TACACCTACTTTTGAACAATGGGGGAGATTAATTGATTAAGCACTTCAAGAGCTTGAAGCC 360  
DB 301 TACACCTACTTTTGAACAATGGGGGAGATTAATTGATTAAGCACTTCAAGAGCTTGAAGCC 360  
QY 361 CGGCAATTTCTATGACACTGAGACATGCAATGACTGTGATGATTAGAACTTGTGTTGAG 420  
DB 361 CGGCAATTTCTATGACACTGAGACATGCAATGACTGTGATGATTAGAACTTGTGTTGAG 420  
QY 421 CCGTGAATGCTGAGACTTGGCCAGCCCTCAGAAAGATTTTAAAGTCAAGAGAGAGCA 480  
DB 421 CCGTGAATGCTGAGACTTGGCCAGCCCTCAGAAAGATTTTAAAGTCAAGAGAGAGCA 480  
QY 481 GAGGAGATTAAGTGGGCACTCCCGGATGATCACTGATCTTGAAGACAGACCTTTG 540  
DB 481 GAGGAGATTAAGTGGGCACTCCCGGATGATCACTGATCTTGAAGACAGACCTTTG 540  
QY 541 AAGTCAGAGCTGTACATATGATCTCAAGTGCAGCTTCTGAGATTGATGATTCAGAA 600  
DB 541 AAGTCAGAGCTGTACATATGATCTCAAGTGCAGCTTCTGAGATTGATGATTCAGAA 600  
QY 601 AGAAAGAGATTTGAGGTTTGAAGCAAAATGACAGTAAACAGCAACCAATTCAAATGTTGA 660  
DB 601 AGAAAGAGATTTGAGGTTTGAAGCAAAATGACAGTAAACAGCAACCAATTCAAATGTTGA 660  
QY 661 ATGAGAGCTTTGAGTCTCACTTACCCGTTGCTGATCCCACTCTCAAGAGCTCTG 720  
DB 661 ATGAGAGCTTTGAGTCTCACTTACCCGTTGCTGATCCCACTCTCAAGAGCTCTG 720  
QY 721 AATATTCCTGATTTACCCCAAGAAATTTAAGATACATCTGAGAGAGTCTTTGAGCAG 780  
DB 721 AATATTCCTGATTTACCCCAAGAAATTTAAGATACATCTGAGAGAGTCTTTGAGCAG 780

QY 781 GAGAAAGCCAGATATCTGTGACTTACAGAGATCCAGTTTTCAGATGCCAATTTCAAG 840  
DB 781 GAGAAAGCCAGATATCTGTGACTTACAGAGATCCAGTTTTCAGATGCCAATTTCAAG 840  
QY 841 GCAATTCACCTTACAGATGATGATCCATTAAGAACCACTCTGCTGTGATATTTGAGACT 900  
DB 841 GCAATTCACCTTACAGATGATGATCCATTAAGAACCACTCTGCTGTGATATTTGAGACT 900  
QY 901 TCAATACAGACTTTTCCATGAGCTGAGATGCTTCAAGGATGATCTGAGCTTCAAGT 960  
DB 901 TCAATACAGACTTTTCCATGAGCTGAGATGCTTCAAGGATGATCTGAGCTTCAAGT 960  
QY 961 GATTCGAGTACCAAGCTTCTCCAAAGACTGACCTTGAAGATTAAGAGAGAGAGAGAG 1020  
DB 961 GATTCGAGTACCAAGCTTCTCCAAAGACTGACCTTGAAGATTAAGAGAGAGAGAGAG 1020  
QY 1021 GTCTTTTGAATAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080  
DB 1021 GTCTTTTGAATAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080  
QY 1081 CCGTGGGAGATGTTCTTCAAGTTCAATTTTACCTGATCTTGAATCCGAGCAATTCCT 1140  
DB 1081 CCGTGGGAGATGTTCTTCAAGTTCAATTTTACCTGATCTTGAATCCGAGCAATTCCT 1140  
QY 1141 AAAAAGCATTTTTCGAGCCCTTGTGAGCTATACAGTGAAGTGTGAAAAGGCGCAG 1200  
DB 1141 AAAAAGCATTTTTCGAGCCCTTGTGAGCTATACAGTGAAGTGTGAAAAGGCGCAG 1200  
QY 1201 CTACAGAGCTGTGAGTAAACAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260  
DB 1201 CTACAGAGCTGTGAGTAAACAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260  
QY 1261 TGTGCTGTTTGTGATCTCTCTGCTGCTTCCCTTCTGAGAGAGAGAGAGAGAG 1320  
DB 1261 TGTGCTGTTTGTGATCTCTCTGCTGCTTCCCTTCTGAGAGAGAGAGAGAGAG 1320  
QY 1321 CTGCTGAGATCTTCTTAACTTCAACCCAGACCAATTCGTTGAGAGAGAGAGAG 1380  
DB 1321 CTGCTGAGATCTTCTTAACTTCAACCCAGACCAATTCGTTGAGAGAGAGAGAG 1380  
QY 1381 TTTCAACCGAG 1440  
DB 1381 TTTCAACCGAG 1440  
QY 1441 AAGAGGTTCTGCGAGAGAGAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500  
DB 1441 AAGAGGTTCTGCGAGAGAGAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500  
QY 1501 CTTGAGCCAAATACATGATCCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560  
DB 1501 CTTGAGCCAAATACATGATCCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560  
QY 1561 TCCATCTCTCTGAG 1620  
DB 1561 TCCATCTCTCTGAG 1620  
QY 1621 ATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
DB 1621 ATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
QY 1677 AAATCAG 1736  
DB 1677 AAATCAG 1736  
QY 1681 AAATCAG 1740  
DB 1681 AAATCAG 1740  
QY 1737 AGGATTAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1796  
DB 1737 AGGATTAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1796  
QY 1741 AGGATTAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800  
DB 1741 AGGATTAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800  
QY 1797 ATCTTAATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1856  
DB 1797 ATCTTAATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1856  
QY 1801 ATCTTAATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860  
DB 1801 ATCTTAATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860  
QY 1857 CAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1916



```

Db      1861  CCGAGCAAGATGATGACAAACAATCTACCTTCATGCGCAGAGGTTGGAGAAATCTC 1920
Qy      1917  CTCAGAGAGAACGGCCATATTTATGTGTGAGATCAAAAGATATGCGCAAGATGTA 1976
Db      1921  CTCAGAGAGAACGGCCATATTTATGTGTGAGATCAAAAGATATGCGCAAGATGTA 1980
Qy      1977  CATGATCCCTTGTGCAATATTAAGCAAGAGTTGAGTTGAAAACTAGAAGCAATG 2036
Db      1981  CATGATCCCTTGTGCAATATTAAGCAAGAGTTGAGTTGAAAACTAGAAGCAATG 2040
Qy      2037  AAAACCTGGCCCTTTAAAAAGAAAGAAAGCTTACCTTCAGATATTTGGTCAATA 2093
Db      2041  AAAACCTGGCCCTTTAAAAAGAAAGAAAGCTTACCTTCAGATATTTGGTCAATA 2097

RESULT 3
US-09-371-347-24
; Sequence 24, Application US/09371347
; Publication No. US20030082676A1
; GENERAL INFORMATION:
; APPLICANT: Roy A. Gravel et al.
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE REDUCTASE:
; TITLE OF INVENTION: CLONING, AND METHODS FOR EVALUATING RISK OF NEURAL TUBE
; TITLE OF INVENTION: DEFECTS, CARDIOVASCULAR DISEASE, AND CANCER
; FILE REFERENCE: 50004/003003
; CURRENT APPLICATION NUMBER: US/09/371.347
; CURRENT FILING DATE: 1999-08-10
; PRIOR APPLICATION NUMBER: 60/071.622
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 09/232.028
; PRIOR FILING DATE: 1999-01-15
; SOFTWARE: PaedSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 3259
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-371-347-24

Query Match      85.7%; Score 1793; DB 10; Length 3259;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2093; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

Qy      1  ATGAGAGGTTTGTGTTACTATATGCTACACAGCAGGAGCAAGGCAAGCCATCGCAGAA 60
Db      80  ATGAGAGGTTTGTGTTACTATATGCTACACAGCAGGAGCAAGGCAAGCCATCGCAGAA 139
Qy      61  GAAATGTGTGAGCAAGCTGTGTGATCATGATTTTCTGCAATCTTCACTGTATTAAGTAA 120
Db      140  GAAATGTGTGAGCAAGCTGTGTGATCATGATTTTCTGCAATCTTCACTGTATTAAGTAA 199
Qy      121  TCGGATTAAGTATGACTTAATAACCGAAACAGCTCCTGTGTGTGTGTGTGTGTGTGTGT 180
Db      200  TCGGATTAAGTATGACTTAATAACCGAAACAGCTCCTGTGTGTGTGTGTGTGTGTGTGT 259
Qy      181  GGACCGGAGAACCCACCGACACAGCCGCGAAAGTTTGTAAAGAAATACAGAAACCAACA 240
Db      260  GGACCGGAGAACCCACCGACACAGCCGCGAAAGTTTGTAAAGAAATACAGAAACCAACA 319
Qy      241  CTGCGCGGTGATTTCTTTGTCTCACTGCGGTATGGGTTTCTGGGTCTCGGTGATTCAGAA 300
Db      320  CTGCGCGGTGATTTCTTTGTCTCACTGCGGTATGGGTTTCTGGGTCTCGGTGATTCAGAA 379
Qy      301  TACACCTACTTTTGCATGGGGGAGATTAATTGAAGAGACTTCAAGAGCTTGGAGCC 360
Db      380  TACACCTACTTTTGCATGGGGGAGATTAATTGAAGAGACTTCAAGAGCTTGGAGCC 439
Qy      361  CGGATTTCTATGACACTGACATGACATGACATGCTGTAGGTTTGAAGACTTGTGTGAG 420
Db      440  CGGATTTCTATGACACTGACATGACATGACATGCTGTAGGTTTGAAGACTTGTGTGAG 499
Qy      421  CCGTGATTTGCTGAGCTCTGCGCAGCCCTCAGAAAGCAATTTTGTAGTCAAGCAGAGCAA 480

```

```

Db      500  CCGTGATTTGCTGAACTTGGCCAGCCCTCAGAAAGCAATTTTAAAGTCAAGCAGAGAGCAA 559
Qy      481  GAGAGATTAAGTGGGCACTCCCGGTGATCATCTGCATCTTGTAGAGACAGACTTGTG 540
Db      560  GAGAGATTAAGTGGGCACTCCCGGTGATCATCTGCATCTTGTAGAGACAGACTTGTG 619
Qy      541  AAGTCAGACTGTCACATTAATGATCTCAAGTGCAGCTTCTGAGATTCATGATTCAGGA 600
Db      620  AAGTCAGACTGTCACATTAATGATCTCAAGTGCAGCTTCTGAGATTCATGATTCAGGA 679
Qy      601  AGAAGAGATTTGAGGTTTGAAGCAAAATGCAAGTGAAGCAAGCAACCAATGATGTA 660
Db      680  AGAAGAGATTTGAGGTTTGAAGCAAAATGCAAGTGAAGCAAGCAACCAATGATGTA 739
Qy      661  ATTGAAGACTTTGAGTCTCACTTACCCGTTGGTATCCCACTCTCAGAGCTCTCTG 720
Db      740  ATTGAAGACTTTGAGTCTCACTTACCCGTTGGTATCCCACTCTCAGAGCTCTCTG 799
Qy      721  AATATTCCTGTTTACCCCGAGAAATTTTACAGTATCATCTGAGAGAGTCTCTGGCCAG 780
Db      800  AATATTCCTGTTTACCCCGAGAAATTTTACAGTATCATCTGAGAGAGTCTCTGGCCAG 859
Qy      781  GAGGAAGCCAGTATCTGTGACTTCAGAGATCCAGTTTTCAGTGCATTTCAAGTCAATTCAAG 840
Db      860  GAGGAAGCCAGTATCTGTGACTTCAGAGATCCAGTTTTCAGTGCATTTCAAGTCAATTCAAG 919
Qy      841  GCAGTTCAACTTACTACGATGATGCGATTAATAACCACTCTGCTGTAGAAATTTGACATT 900
Db      920  GCAGTTCAACTTACTACGATGATGCGATTAATAACCACTCTGCTGTAGAAATTTGACATT 979
Qy      901  TCAATATCAGACTTTTCTATCAGCTGAGATGCTTCAAGCTGATCTGCCCTTAACAGT 960
Db      980  TCAATATCAGACTTTTCTATCAGCTGAGATGCTTCAAGCTGATCTGCCCTTAACAGT 1039
Qy      961  GATTCGAGGTCAAAAGCCCTACCCAGAAATGAGACCTGGAAGATTAAGAGAGCACTGC 1020
Db      1040  GATTCGAGGTCAAAAGCCCTACCCAGAAATGAGACCTGGAAGATTAAGAGAGCACTGC 1099
Qy      1021  GTCCCTTTGAAAAATAAGGACAGACAAAGAAAGAAAGAGTCTTACCCCGCATATTA 1080
Db      1100  GTCCCTTTGAAAAATAAGGACAGACAAAGAAAGAAAGAGTCTTACCCCGCATATTA 1159
Qy      1081  CTTGCGGAGATTTCTCTCAAGTTCAATTTTCTGCTGTGTGTGTGTGTGTGTGTGTGT 1140
Db      1160  CTTGCGGAGATTTCTCTCAAGTTCAATTTTCTGCTGTGTGTGTGTGTGTGTGTGTGT 1219
Qy      1141  AAAAAGCAATTTTGGAGACCCCTGTGACATAACAGTGAACAGTGTGAAAAAGCCAGG 1200
Db      1220  AAAAAGCAATTTTGGAGACCCCTGTGACATAACAGTGAACAGTGTGAAAAAGCCAGG 1279
Qy      1201  CTACAGAGCTGTGACAGTAAACAAGGGGAGCCGATTAATAGCCCTTGTATCAGATGCGC 1260
Db      1280  CTACAGAGCTGTGACAGTAAACAAGGGGAGCCGATTAATAGCCCTTGTATCAGATGCGC 1339
Qy      1261  TGTGCTGTGTTGTGATCTCTCTGCTTCTCTTCTTTCAGGCAACCACTCACTGCTC 1320
Db      1340  TGTGCTGTGTTGTGATCTCTCTGCTTCTCTTCTTTCAGGCAACCACTCACTGCTC 1399
Qy      1321  CTGCTGGAACATTTCTTAACTTCAACCAAGCAATTTCTGTGACAGCTCAAGTTTA 1380
Db      1400  CTGCTGGAACATTTCTTAACTTCAACCAAGCAATTTCTGTGACAGCTCAAGTTTA 1459
Qy      1381  TTTCAACCCAGAAAGCTCAATTTTGTCTTCAACATTTGTGAATTTCTGTCTATGCGACA 1440
Db      1460  TTTCAACCCAGAAAGCTCAATTTTGTCTTCAACATTTGTGAATTTCTGTCTATGCGACA 1519
Qy      1441  ACAAGAGTTCTGCGAAGGAGAGTATGACAGCTGAGCTGAGCTTGTGTGTTGCTTCAATT 1500
Db      1520  ACAAGAGTTCTGCGAAGGAGAGTATGACAGCTGAGCTGAGCTTGTGTGTTGCTTCAATT 1579
Qy      1501  CTTGACGCAACATATCATGATCTCCATGAAAGCAGGGGAAAGCCCTGGGTCTTAAGATA 1560

```

Db 1580 CTTGACCCAAACATCATCATCCCATGAAAGACAGCGGAAAGCCCTGCTCTCAAGATA 1639  
Qy 1561 TCCATCTCTCTCTGGAACAAATCTTCCATTAACAGATGACCCCTCAATCCCATC 1620  
Db 1640 TCCATCTCTCTCTGGAACAAATCTTCCATTAACAGATGACCCCTCAATCCCATC 1699  
Qy 1621 ATAAATGTTGGGTCAGAGAACCGGCAATGCCCCGTTATTTGGGTTCTTCAACAT---AG 1676  
Db 1700 ATAAATGTTGGGTCAGAGAACCGGCAATGCCCCGTTATTTGGGTTCTTCAACATGAGAG 1759  
Qy 1677 AAATCTCCAGAAACAACCCCAATGGAATTTTGAAGCAATGTTGTTTGGCTGC 1736  
Db 1760 AAATCTCCAGAAACAACCCCAATGGAATTTTGAAGCAATGTTGTTTGGCTGC 1819  
Qy 1737 AGGCAATAGGATAGGGAATTAATTAATTAAGGCTCAGACATTTCTTAAAGCATGGG 1796  
Db 1820 AGGCAATAGGATAGGGAATTAATTAATTAAGGCTCAGACATTTCTTAAAGCATGGG 1879  
Qy 1797 ATCTTAATCATCTAAAGGTTCTCTGCAAGATGCTCTGTTGGGAGAGAGAGCC 1856  
Db 1880 ATCTTAATCATCTAAAGGTTCTCTGCAAGATGCTCTGTTGGGAGAGAGAGCC 1939  
Qy 1857 CCAGCAATAGTATGACAGACAAATCATGCTTCATGCGCAGAGGTCGAGAAATCCTC 1916  
Db 1940 CCAGCAATAGTATGACAGACAAATCATGCTTCATGCGCAGAGGTCGAGAAATCCTC 1999  
Qy 1917 CTCGAGAGAAAGGCAATTTATGTTGTTGAGATGCAAAATATGCCCCAAGATGTA 1976  
Db 2000 CTCGAGAGAAAGGCAATTTATGTTGTTGAGATGCAAAATATGCCCCAAGATGTA 2059  
Qy 1977 CATGATGCTCTGTCGAATTAATTAAGCAAGAGTTGAGTTGAAAACTAGAGCAATG 2036  
Db 2060 CATGATGCTCTGTCGAATTAATTAAGCAAGAGTTGAGTTGAAAACTAGAGCAATG 2119  
Qy 2037 AAAACCTGCGCACTTTAAAGAGAAAAACCTTACAGATATTTGTCTATA 2093  
Db 2120 AAAACCTGCGCACTTTAAAGAGAAAAACCTTACAGATATTTGTCTATA 2176

## RESULT 4

US-09-371-347-41  
; Sequence 41, Application US/09371347  
; Publication No. US20030082676A1  
; GENERAL INFORMATION:  
; APPLICANT: Roy A. Gravel et al.  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE REDUCTASE.  
; TITLE OF INVENTION: CLONING, AND METHODS FOR EVALUATING RISK OF NEURAL TUBE  
; FILE REFERENCE: 50004/003003  
; CURRENT APPLICATION NUMBER: US/09/371,347  
; PRIOR FILING DATE: 1999-08-10  
; PRIOR APPLICATION NUMBER: 60/071,622  
; PRIOR FILING DATE: 1998-01-16  
; PRIOR APPLICATION NUMBER: 09/232,028  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 41  
; LENGTH: 2097  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-371-347-41

Query Match 83.2%; Score 1742; DB 10; Length 2097;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2092; Conservative 0; Mismatches 1; Indels 4; Gaps 1;

Qy 1 ATGAGAGAGTTCTGTTATATATGCTACAGCAGGAGCAGGCAAGGCCATCGAGAA 60  
Db 1 ATGAGAGAGTTCTGTTATATATGCTACAGCAGGAGCAGGCAAGGCCATCGAGAA 60  
Qy 61 GAAATGTGACAGACTGTGATCAATGATTTTCTGAGATCTTCACTGATTAAGTAA 120  
Db 61 GAAATGTGACAGACTGTGATCAATGATTTTCTGAGATCTTCACTGATTAAGTAA 120

Db 61 GAAATGTGACAGACTGTGATCAATGATTTTCTGAGATCTTCACTGATTAAGTAA 120  
Qy 121 TCCGATTAAGTATGACCTAATAAACCGAAACAGCTCCTGTTGTTGTTGTTCTACAG 180  
Db 121 TCCGATTAAGTATGACCTAATAAACCGAAACAGCTCCTGTTGTTGTTGTTCTACAG 180  
Qy 181 GGCACCGGAGACCCACCGACACAGCCCGCAAGTTGTTAAGAAATACAGAACCAACA 240  
Db 181 GGCACCGGAGACCCACCGACACAGCCCGCAAGTTGTTAAGAAATACAGAACCAACA 240  
Qy 241 CTGCGGATGATTTCTTGTCTCACTGCGGATATGAGTTATCTGAGTTCTCGGTATTCAGA 300  
Db 241 CTGCGGATGATTTCTTGTCTCACTGCGGATATGAGTTATCTGAGTTCTCGGTATTCAGA 300  
Qy 301 TACACCTACTTTTGAATGGGGGAAAGATTAATGATTAACAGACTTCAAGAGCTTGAAGCC 360  
Db 301 TACACCTACTTTTGAATGGGGGAAAGATTAATGATTAACAGACTTCAAGAGCTTGAAGCC 360  
Qy 361 CCGCATTTCTATGACACTGACATGACATGATGATGATGATTTAGAACTTGATGAG 420  
Db 361 CCGCATTTCTATGACACTGACATGACATGATGATGATGATTTAGAACTTGATGAG 420  
Qy 421 CCGTGAATGCTGACCTGCGCAGGCTCAGAAAGCAATTTAGTCAAGAGAGACA 480  
Db 421 CCGTGAATGCTGACCTGCGCAGGCTCAGAAAGCAATTTAGTCAAGAGAGACA 480  
Qy 481 GAGAGATTAAGTGGGCACTCCCGGTGATACCTGATCTTGAAGGACAGACCTTGAG 540  
Db 481 GAGAGATTAAGTGGGCACTCCCGGTGATACCTGATCTTGAAGGACAGACCTTGAG 540  
Qy 541 AAGTCAAGCTGCTACATTAATCTCAATGAGCTTGAAGTTGAGATTCAGATTCAGAGA 600  
Db 541 AAGTCAAGCTGCTACATTAATCTCAATGAGCTTGAAGTTGAGATTCAGATTCAGAGA 600  
Qy 601 AGAAAGATTCGAGGTTTGAAGCAAAATGCAAGTGAACAGACCAATTCATGTTGA 660  
Db 601 AGAAAGATTCGAGGTTTGAAGCAAAATGCAAGTGAACAGACCAATTCATGTTGA 660  
Qy 661 ATTGAAGCTTGAAGTCTCACTTACCCGTTGGTATACCCCACTGCAAGGCTCTCG 720  
Db 661 ATTGAAGCTTGAAGTCTCACTTACCCGTTGGTATACCCCACTGCAAGGCTCTCG 720  
Qy 721 AATATTCGTTTACCCCGAATATTTACAGTACATCTGACAGAGTCTTGGCCAG 780  
Db 721 AATATTCGTTTACCCCGAATATTTACAGTACATCTGACAGAGTCTTGGCCAG 780  
Qy 781 GAGGAAGCAAGTATCTGTGACTTCAAGAGATTCAGATTTCAAGTGCATATTCAGAG 840  
Db 781 GAGGAAGCAAGTATCTGTGACTTCAAGAGATTCAGATTTCAAGTGCATATTCAGAG 840  
Qy 841 GCAATTCACATTAATGCAATGATGCAATTAACCACTGCTGCTGTAATTTGACATT 900  
Db 841 GCAATTCACATTAATGCAATGATGCAATTAACCACTGCTGCTGTAATTTGACATT 900  
Qy 901 TCAATTAACAGACTTTTCTATCAAGCTGAGATGCTTCAAGGTTATCTGCTTAACAT 960  
Db 901 TCAATTAACAGACTTTTCTATCAAGCTGAGATGCTTCAAGGTTATCTGCTTAACAT 960  
Qy 961 GATTCGAGTACCAAGCTTCTCCAAAGATGAGCTTGAAGATTAAGAGACATGCTC 1020  
Db 961 GATTCGAGTACCAAGCTTCTCCAAAGATGAGCTTGAAGATTAAGAGACATGCTC 1020  
Qy 1021 GTTCCTTTGAAATTAAGGACACCAAGAAAGAGCTTACCTTACCAGCATTA 1080  
Db 1021 GTTCCTTTGAAATTAAGGACACCAAGAAAGAGCTTACCTTACCAGCATTA 1080  
Qy 1081 CCGCGGAGATGTTCTCTCAAGTTCAATTTTACCTGAGTCTTGAATCCGAGCAATTCCT 1140  
Db 1081 CCGCGGAGATGTTCTCTCAAGTTCAATTTTACCTGAGTCTTGAATCCGAGCAATTCCT 1140  
Qy 1141 AAAAAGCAATTTTTCAGAGCCCTTGTGACATTAACAGTGAAGTGTGAAAAGCGCAGG 1200  
Db 1141 AAAAAGCAATTTTTCAGAGCCCTTGTGACATTAACAGTGAAGTGTGAAAAGCGCAGG 1200

1201 CTACAGAGCTGTGAGTAAACAAGGGGACGCCGATTAAGCCGTTTGTACAGATGCC 1260  
1201 CTACAGAGCTGTGAGTAAACAAGGGGACGCCGATTAAGCCGTTTGTACAGATGCC 1260  
1261 TGTGCTGTGTTGGATCTCTCTCGCTTTCCCTTTCTTGCCAGCACAATCACTACTC 1320  
1261 TGTGCTGTGTTGGATCTCTCTCGCTTTCCCTTTCTTGCCAGCACAATCACTACTC 1320  
1261 TGTGCTGTGTTGGATCTCTCTCGCTTTCCCTTTCTTGCCAGCACAATCACTACTC 1320  
1321 CTGCTGAAATCTTCTTAACTTCAACCCAGACATATTCGTCAGCTCAAGTTTA 1380  
1321 CTGCTGAAATCTTCTTAACTTCAACCCAGACATATTCGTCAGCTCAAGTTTA 1380  
1381 TTTCACCCAGAAAGCTCCATTTTGTCTTCAACATTTGTGGAATTTCTTACTGTCACA 1440  
1381 TTTCACCCAGAAAGCTCCATTTTGTCTTCAACATTTGTGGAATTTCTTACTGTCACA 1440  
1441 ACAGAGTTCTGCGGAAAGGAGATGATACAGGCTGCTGCTTTGTTGTTGCTTCACTT 1500  
1441 ACAGAGTTCTGCGGAAAGGAGATGATACAGGCTGCTGCTTTGTTGTTGCTTCACTT 1500  
1501 CTTCAGCAAAATATATGATGATCCATGAAAGAGAGGAGGAGGAGGAGGAGGAGGAGG 1560  
1501 CTTCAGCAAAATATATGATGATCCATGAAAGAGAGGAGGAGGAGGAGGAGGAGGAGG 1560  
1561 TCCATCTCTCTGGAACAACAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
1561 TCCATCTCTCTGGAACAACAATTTCTTCACTTACAGATGACCCCTCAATCCCATC 1620  
1621 ATATAGTGGGTGCGAGAACCGGATAGCCCGTTTATGGGTTCTTCAACAATAGAGAG 1680  
1621 ATATAGTGGGTGCGAGAACCGGATAGCCCGTTTATGGGTTCTTCAACAATAGAGAG 1680  
1677 AAATCTCAAGAAACAACCCAGATGAAATTTTGGAGCAATGTGTGTTTGTGTTGGCTGC 1736  
1681 AAATCTCAAGAAACAACCCAGATGAAATTTTGGAGCAATGTGTGTTTGTGTTGGCTGC 1740  
1737 AGGCAATAGGATAGGATTAATCTATTCAGAAAAGAGCTCAGACATTTCTTAAAGATGG 1796  
1741 AGGCAATAGGATAGGATTAATCTATTCAGAAAAGAGCTCAGACATTTCTTAAAGATGG 1800  
1797 ATCTTAATCTATTTAAAGGTTTCTTCTCAAGAGATCTCTGTTGGGAGAGAGAGGCC 1856  
1801 ATCTTAATCTATTTAAAGGTTTCTTCTCAAGAGATCTCTGTTGGGAGAGAGAGGCC 1860  
1857 CCAGCAAGATGATGACAAGACAATCAGCTTCAATGAGGAGGAGGAGGAGGAGGAGGAG 1916  
1861 CCAGCAAGATGATGACAAGACAATCAGCTTCAATGAGGAGGAGGAGGAGGAGGAGGAG 1920  
1917 CTCCAGAGAAAGGAGGATTAATTTATGTGTGAGATGCAAAAGATATGAGCAAGATGTA 1976  
1921 CTCCAGAGAAAGGAGGATTAATTTATGTGTGAGATGCAAAAGATATGAGCAAGATGTA 1980  
1977 CATGATGAGCTTTGTGCAAAATATATGAGCAAGAGGTTGAGTTGAAAAATCTAGAGCAATG 2036  
1981 CATGATGAGCTTTGTGCAAAATATATGAGCAAGAGGTTGAGTTGAAAAATCTAGAGCAATG 2040  
2037 AAAAAGCTGAGCACTTTAAAGAAAGAAAGAAAGCTTCAAGATATTTGTCATTA 2093  
2041 AAAAAGCTGAGCACTTTAAAGAAAGAAAGAAAGCTTCAAGATATTTGTCATTA 2097

RESULT 5  
US-09-371-347-43  
; Sequence 43: Application US/09371347  
; Publication No. US20030082676A1  
; GENERAL INFORMATION:  
; APPLICANT: Roy A. Gravel et al.  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE REDUCTASE;  
; TITLE OF INVENTION: CLONING, AND METHODS FOR EVALUATING RISK OF NEURAL TUBE  
; TITLE OF INVENTION: DEFECTS, CARDIOVASCULAR DISEASE, AND CANCER  
; FILE REFERENCE: 50004/003003  
; CURRENT APPLICATION NUMBER: US/09/371.347

; CURRENT FILING DATE: 1999-08-10  
; PRIOR APPLICATION NUMBER: 60/071,622  
; PRIOR FILING DATE: 1998-01-16  
; PRIOR APPLICATION NUMBER: 09/232,028  
; PRIOR FILING DATE: 1999-01-15  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 43  
; LENGTH: 2097  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-371-347-43

Query Match 83.2%; Score 1742; DB 10; Length 2097;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2092; Conservative 0; Mismatches 1; Indels 4; Gaps 1;

1 ATGAGAGGTTTCTGTACTATATGCTTACACAGCAGGAGCAGGCAAGCCATCCGAGAA 60  
1 ATGAGAGGTTTCTGTACTATATGCTTACACAGCAGGAGCAGGCAAGCCATCCGAGAA 60  
61 GAAATGTGAGCAGCTGTGATGATGATTTTCTGAGATCTTCACTGATATGTA 120  
61 GAAATGTGAGCAGCTGTGATGATGATTTTCTGAGATCTTCACTGATATGTA 120  
121 TCCGATATGATGACCTTAAACCCGAAACAGCTCTTGTGTGTGTTGTTTACAG 180  
121 TCCGATATGATGACCTTAAACCCGAAACAGCTCTTGTGTGTGTTGTTTACAG 180  
181 GGCACCGGAGACCCACCCGACACAGCCGCGAAGTTGTTAAGAAATACAGAACCA 240  
181 GGCACCGGAGACCCACCCGACACAGCCGCGAAGTTGTTAAGAAATACAGAACCA 240  
241 CTGCGGTTGATTTCTTGTCTCACTGCGGATAGGTTACTGGGCTCGGTTAGAA 300  
241 CTGCGGTTGATTTCTTGTCTCACTGCGGATAGGTTACTGGGCTCGGTTAGAA 300  
301 TACACCTATTTTGAATGAGGAGGAGATTAATGATTAACGACTTCAAGAGCTTGA 360  
301 TACACCTATTTTGAATGAGGAGGAGATTAATGATTAACGACTTCAAGAGCTTGA 360  
361 CGGATTTCTATGACCTGACATGAGATGATGATGATGATGATGATGATGATGATG 420  
361 CGGATTTCTATGACCTGACATGAGATGATGATGATGATGATGATGATGATGATG 420  
421 CGGTGATGCTGAGCTGCGCAGCCCTCAGAAACATTTTATGATCAAGAGAGCA 480  
421 CGGTGATGCTGAGCTGCGCAGCCCTCAGAAACATTTTATGATCAAGAGAGCA 480  
481 GAGAGATATGAGGAGCACTCCGATGAGATCACTGATCTTGAAGAGAGAGAGAG 540  
481 GAGAGATATGAGGAGCACTCCGATGAGATCACTGATCTTGAAGAGAGAGAGAG 540  
541 AAGTCAAGCTGTACACATGATGATGATGATGATGATGATGATGATGATGATG 600  
541 AAGTCAAGCTGTGTACACATGATGATGATGATGATGATGATGATGATGATGATG 600  
601 AGAAAGATTTCTGAGGTTTGAAGCAAAATGAGTGAACAGCAACCAATCCATGTTGA 660  
601 AGAAAGATTTCTGAGGTTTGAAGCAAAATGAGTGAACAGCAACCAATCCATGTTGA 660  
661 ATTGAAGCTTTGAGTCTTCACTTACCCGTTGGATACCCCACTCTCAAGGCTCTG 720  
661 ATTGAAGCTTTGAGTCTTCACTTACCCGTTGGATACCCCACTCTCAAGGCTCTG 720  
721 AATATTCCTGTTTACCCCAAGATATTTTACAGATATCTGAGAGAGTCTTGGGCA 780  
721 AATATTCCTGTTTACCCCAAGATATTTTACAGATATCTGAGAGAGTCTTGGGCA 780  
781 GAGGAAGCAAGATATCTGATGATGATGATGATGATGATGATGATGATGATGATG 840  
781 GAGGAAGCAAGATATCTGATGATGATGATGATGATGATGATGATGATGATGATG 840

Oy	841	GCAGTTCAACTTACCTACGAATGATGCCATAAAAAACA	CTGCGTGGTGAATGGAACATT	900
Db	841	GCAGTTCAACTTACCTACGAATGATGCCATAAAAAACA	CTGCGTGGTGAATGGAACATT	900
Oy	901	TCAATAACAGACTTTCCTATCAGCTTGAGATGCTTCA	CGCGTGAATCTGCGCTTAAAGT	960
Db	901	TCAATAACAGACTTTCCTATCAGCTTGAGATGCTTCA	CGCGTGAATCTGCGCTTAAAGT	960
Oy	961	GATTCTGAGGTACAAAGCTTACTCCAAAGACTGACG	CTTGAAGATTAAGAGACACTGC	1020
Db	961	GATTCTGAGGTACAAAGCTTACTCCAAAGACTGACG	CTTGAAGATTAAGAGACACTGC	1020
Oy	1021	GTCCCTTTGAAAATAAAGCAGACACAAGAAAGAAAG	AGGACTTACCTTACCCACATATA	1080
Db	1021	GTCCCTTTGAAAATAAAGCAGACACAAGAAAGAAAG	AGGACTTACCTTACCCACATATA	1080
Oy	1081	CCTCGGGGATGTTCTCTCAGTTCATTTTACCTGATG	CTTGAATTCGAGCAATTCCT	1140
Db	1081	CCTCGGGGATGTTCTCTCAGTTCATTTTACCTGATG	CTTGAATTCGAGCAATTCCT	1140
Oy	1141	AAAAAGGCAATTTTTCGAGCCCTTGATGACTATA	CCAGTGAACAGTGTGAAGACGACG	1200
Db	1141	AAAAAGGCAATTTTTCGAGCCCTTGATGACTATA	CCAGTGAACAGTGTGAAGACGACG	1200
Oy	1201	CTACAGAGCTGTGCAATTAACAAGGGCAGCCGATTA	TAGCCGCTTTGTATCGAGATGCC	1260
Db	1201	CTACAGAGCTGTGCAATTAACAAGGGCAGCCGATTA	TAGCCGCTTTGTATCGAGATGCC	1260
Oy	1261	TGTGCGTCTGTTGGATCTGCTCGCTTCCCTTTCG	ACAGCCACGACTAGTCTC	1320
Db	1261	TGTGCGTCTGTTGGATCTGCTCGCTTCCCTTTCG	ACAGCCACGACTAGTCTC	1320
Oy	1321	CTGCTCGAACATCTTCTTAACTTCAACCCAGACCAT	ATTGATGTGACAGCTCAAGTTTA	1380
Db	1321	CTGCTCGAACATCTTCTTAACTTCAACCCAGACCAT	ATTGATGTGACAGCTCAAGTTTA	1380
Oy	1381	TTTTCACCCAGAAAAGCTCCATTTTGTCTTCAACAT	TGTGGAAATTTCTGTCTACGCCACA	1440
Db	1381	TTTTCACCCAGAAAAGCTCCATTTTGTCTTCAACAT	TGTGGAAATTTCTGTCTACGCCACA	1440
Oy	1441	ACAGAGGTCTGCGGAGAGGGATATGATCAGGCTGG	CGGCGCTTGGTGTGCTTCAAGTT	1500
Db	1441	ACAGAGGTCTGCGGAGAGGGATATGATCAGGCTGG	CGGCGCTTGGTGTGCTTCAAGTT	1500
Oy	1501	CTTACAGCCAAACATACATGATCCCATGAAACAG	ACGGGGAAAGCCCTTGACTTAAAGTA	1560
Db	1501	CTTACAGCCAAACATACATGATCCCATGAAACAG	ACGGGGAAAGCCCTTGACTTAAAGTA	1560
Oy	1561	TCGATCTCTCTCGAACAAACAATTCCTTGCATTTA	CCAGATGACCCCTCAATCCCATC	1620
Db	1561	TCGATCTCTCTCGAACAAACAATTCCTTGCATTTA	CCAGATGACCCCTCAATCCCATC	1620
Oy	1621	ATAATGTGTGGGTCCAGGAACCGGCATAGCCCCG	TTATTGGGTTCTTACACAT----	1676
Db	1621	ATAATGTGTGGGTCCAGGAACCGGCATAGCCCCG	TTATTGGGTTCTTACACAT----	1676
Oy	1677	AAACTCCAAAGAACAAACCCAGATGGAATTTTGG	AGATGATGTTTGGACTGC	1736
Db	1677	AAACTCCAAAGAACAAACCCAGATGGAATTTTGG	AGATGATGTTTGGACTGC	1736
Oy	1737	AGGCAATAGATAGGAGTTATCTATTCCGAAAAGAG	CTCAGACATTTCTTAAAGCATGGG	1796
Db	1737	AGGCAATAGATAGGAGTTATCTATTCCGAAAAGAG	CTCAGACATTTCTTAAAGCATGGG	1796
Oy	1797	ATCTTAACTCATCTTAAAGGTTTCTTCTTCAAGAG	ATGCTCTCTGTTGGGAGAGAAAGCC	1860
Db	1797	ATCTTAACTCATCTTAAAGGTTTCTTCTTCAAGAG	ATGCTCTCTGTTGGGAGAGAAAGCC	1860
Oy	1857	CCAGCAAGATATGTACAGAACAAATCAGGCTTCA	TGCGACAGAGTGGCCAGATCTCTC	1916
Db	1857	CCAGCAAGATATGTACAGAACAAATCAGGCTTCA	TGCGACAGAGTGGCCAGATCTCTC	1916
Oy	1917	CTCCAGAGAAACGGCATATTTATGTGTGAGAGATG	CAAAAGATATGGCCAGAGATGTA	1976
Db	1917	CTCCAGAGAAACGGCATATTTATGTGTGAGAGATG	CAAAAGATATGGCCAGAGATGTA	1976

Dd	1921	CTCCAGGAAACGGCCATATTTAATGTGTGGAATGCMAAAGAATATGSCCAAGATGTA	1980
Qy	1977	CATGATGCCCTTGTTCGAATATAATAGCAAGAGGTTGAGTTGAAAACTAGAACAAATG	2036
Dd	1981	CATGATGCCCTTGTTCGAATATAATAGCAAGAGGTTGAGTTGAAAACTAGAACAAATG	2040
Qy	2037	AAAACCTTGCGCATCTTTAAAAAGAAAACGCTACTTCACGATATTTGGTCTATA	2093
Dd	2041	AAAACCTTGCGCATCTTTAAAAAGAAAACGCTACTTCAGATATTTGGTCTATA	2097
 RESULT 6 US-09-371-347-45 ; Sequence 45, Application US/09371347 ; Publication No. US20030082676n1 ; GENERAL INFORMATION: ; APPLICANT: Roy A. Grave et al. ; TITLE OF INVENTION: HUMAN METHYLONINE SYNTHASE REDUCTASE; ; TITLE OF INVENTION: CLONING, AND METHODS FOR EVALUATING RISK OF NEURAL TUBE ; FILE REFERENCE: 50004/003003 ; CURRENT APPLICATION NUMBER: US/09/371,347 ; PRIOR FILING DATE: 1999-08-10 ; PRIOR APPLICATION NUMBER: 60/071,622 ; PRIOR FILING DATE: 1998-01-16 ; PRIOR APPLICATION NUMBER: 09/232,028 ; PRIOR FILING DATE: 1999-01-15 ; NUMBER OF SEQ ID NOS: 51 ; SOFTWARE: PasteSeq for Windows Version 4.0 ; SEQ ID NO 45 ; LENGTH: 2094 ; TYPE: DNA ; ORGANISM: Homo sapiens US-09-371-347-45			
 Query Match      80.1%; Score 1677; DB 10; Length 2094; Best Local Similarity    100.0%; Pred. No. 0; Matches 1677; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	AATGAGGAGGCTTCTGTTACTATATGCTACACAGACAGGAGCAGGCAAGGCCATGSCAGAA	60
Dd	1	AATGAGGAGGCTTCTGTTACTATATGCTACACAGACAGGAGCAGGCAAGGCCATGSCAGAA	60
Qy	61	GAAATGTGTGAGCAAGCTGTGTATCATAGATTTTCTGCAGATCTTCACTGTATTAAGTAA	120
Dd	61	GAAATGTGTGAGCAAGCTGTGTATCATAGATTTTCTGCAGATCTTCACTGTATTAAGTAA	120
Qy	121	TCCGATAATGATGACCTTAAAAACCGAAACAGCTCCTCTTGTGTGTGTGTTCTACACAG	180
Dd	121	TCCGATAATGATGACCTTAAAAACCGAAACAGCTCCTCTTGTGTGTGTGTTCTACACAG	180
Qy	181	GGCACCGGAGACCCACCCGACACACAGCCCGCAAGTTTGTTAAGGAATACAAACCAAACA	240
Dd	181	GGCACCGGAGACCCACCCGACACACAGCCCGCAAGTTTGTTAAGGAATACAAACCAAACA	240
Qy	241	CTGCGCGGTGATTTCTTTTGTCTCACCTGCGGATGAGGTTACTGAGGCTCGGTGATTCAGAA	300
Dd	241	CTGCGCGGTGATTTCTTTTGTCTCACCTGCGGATGAGGTTACTGAGGCTCGGTGATTCAGAA	300
Qy	301	TACACCTTACTTTTGCAATGGGGGGAGATTAATGATAAACAATTCAGAGCTTGAAGCC	360
Dd	301	TACACCTTACTTTTGCAATGGGGGGAGATTAATGATAAACAATTCAGAGCTTGAAGCC	360
Qy	361	CGGCAATTTCTATGACATGACATGACAGATGACCTGTGTAAGTTTAGAATCTGTGTGAG	420
Dd	361	CGGCAATTTCTATGACATGACATGACAGATGACCTGTGTAAGTTTAGAATCTGTGTGAG	420
Qy	421	CCGTGATATTTGCTGACACTCTGGCCAGACCCTCAGAAAGCAATTTTAGTCAAGCAGAGACA	480
Dd	421	CCGTGATATTTGCTGACACTCTGGCCAGACCCTCAGAAAGCAATTTTAGTCAAGCAGAGACA	480
Qy	481	GAGGAGATTAATGAGGCGACTCCGGGTGGCATACCTGCATCTTGAAGACAGACCTGTGT	540

Db 481 GAGGAAATTAAGTGGCGACTCCGGTGGCATCACTGGCACTTGGAGACAGACTTGG 540  
Qy 541 AAGTCAGAGCTGTACACATTTGATCTCAAGTCAGAGCTTCTGATTCGATGATTCAGGA 600  
Db 541 AAGTCAGAGCTGTACACATTTGATCTCAAGTCAGAGCTTCTGATTCGATGATTCAGGA 600  
Qy 601 AGAAGGATTCGAGCTTTGAAAGCAAAATGAGTGAACAGCAACCAATCCATGTTGA 660  
Db 601 AGAAGGATTCGAGCTTTGAAAGCAAAATGAGTGAACAGCAACCAATCCATGTTGA 660  
Qy 661 ATGAAGACTTGAAGCTCCTACCTTACCCGTTGGTACCCCACTCTCAAGGCTCTGG 720  
Db 661 ATGAAGACTTGAAGCTCCTACCTTACCCGTTGGTACCCCACTCTCAAGGCTCTGG 720  
Qy 721 AATATTCCTGGTTTACCCCGAATATTTTACAGGTACATCTGAGAGAGTCTTGGCCAG 780  
Db 721 AATATTCCTGGTTTACCCCGAATATTTTACAGGTACATCTGAGAGAGTCTTGGCCAG 780  
Qy 781 GAGGAAAGCCAGATTCGTGACTTCAGCAGATTCAGTCTTCAAGTCCAAATTTCAAG 840  
Db 781 GAGGAAAGCCAGATTCGTGACTTCAGCAGATTCAGTCTTCAAGTCCAAATTTCAAG 840  
Qy 841 GCAGTCAACTTACTCGAATGATGCAATTAACAACCTCTGCTGGTGAATTCAGAT 900  
Db 841 GCAGTCAACTTACTCGAATGATGCAATTAACAACCTCTGCTGGTGAATTCAGAT 900  
Qy 901 TCAATATCAGACTTTTCTATCAGACCTGAGATGCTTCAAGCTTCAAGCTTCAAGT 960  
Db 901 TCAATATCAGACTTTTCTATCAGACCTGAGATGCTTCAAGCTTCAAGCTTCAAGT 960  
Qy 961 GATTCGAGGTACAAAGCTTCTCAAGACTGCAAGCTGGAAGATTAAGAGAGCACTGC 1020  
Db 961 GATTCGAGGTACAAAGCTTCTCAAGACTGCAAGCTGGAAGATTAAGAGAGCACTGC 1020  
Qy 1021 GTCTTTTGAATTAAGAGAGACACAAAGAGAGAGTACTTACCCGACATTA 1080  
Db 1021 GTCTTTTGAATTAAGAGAGACACAAAGAGAGAGTACTTACCCGACATTA 1080  
Qy 1081 CCTGCGGAGATGTTCTCTCAAGTCAATTTTACCTGCTGTGAATTCGAGCAATTTCT 1140  
Db 1081 CCTGCGGAGATGTTCTCTCAAGTCAATTTTACCTGCTGTGAATTCGAGCAATTTCT 1140  
Qy 1141 AAAAAGGCAATTTTGGAGAGCCCTTGTGATATACAGTGAACAGTGTGAAAAGCCGAG 1200  
Db 1141 AAAAAGGCAATTTTGGAGAGCCCTTGTGATATACAGTGAACAGTGTGAAAAGCCGAG 1200  
Qy 1201 CTACAGAGGTGAGAGTAAACAAGGGGAGCCGATATATAGCCGCTTGAAGAGATGCC 1260  
Db 1201 CTACAGAGGTGAGAGTAAACAAGGGGAGCCGATATATAGCCGCTTGAAGAGATGCC 1260  
Qy 1261 TGTGCTGCTGTGTGATCTCTCTCAAGTCAATTTTACCTGCTGTGAATTCGAGCAAT 1320  
Db 1261 TGTGCTGCTGTGTGATCTCTCTCAAGTCAATTTTACCTGCTGTGAATTCGAGCAAT 1320  
Qy 1321 CTGCTGCAATCTTCTTAACTTCAACCAAGACCAATATTCGTGTGAAGCTCAAGTTA 1380  
Db 1321 CTGCTGCAATCTTCTTAACTTCAACCAAGACCAATATTCGTGTGAAGCTCAAGTTA 1380  
Qy 1381 TTTTCAACCAAGAAAGCTCCATTTTGTCTGAACAATGTGAAATTTCTGTCTACGCA 1440  
Db 1381 TTTTCAACCAAGAAAGCTCCATTTTGTCTGAACAATGTGAAATTTCTGTCTACGCA 1440  
Qy 1441 ACAGAGGTTCTGCGAAGAGATATGACAGGCTGCGCTTGTGTGTTGCTTCAAGT 1500  
Db 1441 ACAGAGGTTCTGCGAAGAGATATGACAGGCTGCGCTTGTGTGTTGCTTCAAGT 1500  
Qy 1501 CTTCAGGCAAAATACATGATCCCATGAAGAAGCCGAGAAAGCCCTGAGCTCTTAAGATA 1560  
Db 1501 CTTCAGGCAAAATACATGATCCCATGAAGAAGCCCTGAGCTCTTAAGATA 1560  
Qy 1561 TCCATCTCTCTCGAAGAAATTTCTTTCACATTAAGATGAGCCCTCAATCCCATC 1620  
Db 1561 TCCATCTCTCTCGAAGAAATTTCTTTCACATTAAGATGAGCCCTCAATCCCATC 1620

Db 1561 TCCATCTCTCTCGAAGAAATTTCTTTCACATTAAGATGAGCCCTCAATCCCATC 1620  
Qy 1621 AATATGAGGTTCCAGAAACCGGATAGCCCGGTTATTTGGGTTCTTACATATAGA 1677  
Db 1621 AATATGAGGTTCCAGAAACCGGATAGCCCGGTTATTTGGGTTCTTACATATAGA 1677  
RESULT 7  
US-10-741-600-692  
; Sequence 692, Application US/10741600  
; Publication No. US20050026169A1  
; GENERAL INFORMATION:  
; APPLICANT: CARILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; FILE REFERENCE: CL001499  
; CURRENT APPLICATION NUMBER: US/10/741,600  
; NUMBER OF SEQ ID NOS: 73997  
; SOFTWARE: PatsSeq for Windows Version 4.0  
; SEQ ID NO 692  
; LENGTH: 3256  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-741-600-692  
Query Match 42.0%; Score 879; DB 21; Length 3256;  
Best Local Similarity 99.1%; Pred. No. 0;  
Matches 1579; Conservative 0; Mismatches 14; Indels 0; Gaps 0;  
Qy 67 TGTGAGCAGCTGTGTATCATGATTTTCTGAGATCTTCACTGTATGATTCGAT 126  
Db 160 TGTGAGCAGCTGTGTATCATGATTTTCTGAGATCTTCACTGTATGATTCGAT 219  
Qy 127 AAGTATGACTTAAACCGAAACAGCTCTTGTGTGTGTGTTCTACACGGGCAAC 186  
Db 220 AAGTATGACTTAAACCGAAACAGCTCTTGTGTGTGTGTTCTACACGGGCAAC 279  
Qy 187 GAGAGCCACCCGACACAGCCCGCAAGTTGTTAAGAAATACAAACCAACATGCGG 246  
Db 280 GAGAGCCACCCGACACAGCCCGCAAGTTGTTAAGAAATACAAACCAACATGCGG 339  
Qy 247 GTTGAATTCCTTGTCTACCTGCGGTATGGTTATCGGGTCTGGGTATTCAGATTCACC 306  
Db 340 GTTGAATTCCTTGTCTACCTGCGGTATGGTTATCGGGTCTGGGTATTCAGATTCACC 399  
Qy 307 TACTTTGCAATGGGGGGAATTAATTGATTAACGACTTCAGAGCTTGAAGCCCGCAT 366  
Db 400 TACTTTGCAATGGGGGGAATTAATTGATTAACGACTTCAGAGCTTGAAGCCCGCAT 459  
Qy 367 TTCTATGACATGACATGACATGACATGATGATGATGATGATGATGATGATGATGAT 426  
Db 460 TTCTATGACATGACATGACATGACATGATGATGATGATGATGATGATGATGATGAT 519  
Qy 427 ATGTGAGACTGTGGGAGAGCCCTCAGAAAGCAATTTTATAGTCAAGAGAGCAAGAGAG 486  
Db 520 ATGTGAGACTGTGGGAGAGCCCTCAGAAAGCAATTTTATAGTCAAGAGAGCAAGAGAG 579  
Qy 487 ATTAAGTGGCACTCCCGGTGGCATCACTGCACTCTTGAAGAGCAAGCTTGAAGTCA 546  
Db 580 ATTAAGTGGCACTCCCGGTGGCATCACTGCACTCTTGAAGAGCAAGCTTGAAGTCA 639  
Qy 547 GAGCTCTACATTTGAATCTCAAGTCAAGCTTCTGAGATTCGATGATTCAGAGAGAG 606  
Db 640 GAGCTCTACATTTGAATCTCAAGTCAAGCTTCTGAGATTCGATGATTCAGAGAGAG 659  
Qy 607 GATTCGAGGTTTGAAGCAAAATGAGTGAACAGCAACCAATCCAAATGTTGAATTGAA 666  
Db 700 GATTCGAGGTTTGAAGCAAAATGAGTGAACAGCAACCAATCCAAATGTTGAATTGAA 759  
Qy 667 GACTTGAAGCTCCTTACCCGTTGGTACCCCACTCTCAAGAGCTCTGAATATT 726  
Db 760 GACTTGAAGCTCCTTACCCGTTGGTACCCCACTCTCAAGAGCTCTGAATATT 819

QY 727 CCTGTTTACCCCGAATATTTCAGATACATCTGACAGAGTCTCTTGCCGACGAGAA 786  
DB 820 CCGGTTTACCCCGAATATTTCAGATACATCTGACAGAGTCTCTTGCCGACGAGAA 879  
QY 787 ACCCAAGTATCTGTGATCTGACAGATCCAGTTTTCAGAGTCCAAATTTCAAGGCAATT 846  
DB 880 ACCCAAGTATCTGTGATCTGACAGATCCAGTTTTCAGAGTCCAAATTTCAAGGCAATT 939  
QY 847 CAATCTACAGATGATGCAATTAACACATCTGCTGTGATGCAATTTCAAT 906  
DB 940 CAATCTACAGATGATGCAATTAACACACATCTGCTGTGATGCAATTTCAAT 999  
QY 907 ACAGACTTTTCCATCAGCTGAGAGATGCTTCAAGGATCTGCTTCAAGTATCT 966  
DB 1000 ACAGACTTTTCCATCAGCTGAGAGATGCTTCAAGGATCTGCTTCAAGTATCT 1059  
QY 967 GAGTACCAAGCTTCTCAAGAGCTGCAAGCTTGAAGATTAAGAGAGACATGCTCTT 1026  
DB 1060 GAGTACCAAGCTTCTCAAGAGCTGCAAGCTTGAAGATTAAGAGAGACATGCTCTT 1119  
QY 1027 TTGAATAATTAAGGACGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1086  
DB 1120 TTGAATAATTAAGGACGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1179  
QY 1087 GAGTGTCTCTCAGATCTTCTTCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 1146  
DB 1180 GAGTGTCTCTCAGATCTTCTTCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 1239  
QY 1147 GCAATTTTTCAGAGCTTGTGATCTTACAGTGAAGTGTGAGAGAGAGAGAGAGAG 1206  
DB 1240 GCAATTTTTCAGAGCTTGTGATCTTACAGTGAAGTGTGAGAGAGAGAGAGAGAG 1299  
QY 1207 GAGCTGTGAGTAAACAAGGACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1266  
DB 1300 GAGCTGTGAGTAAACAAGGACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1359  
QY 1267 TGCCTGTGAGTCT 1326  
DB 1360 TGCCTGTGAGTCT 1419  
QY 1327 GAAATCTTCTCTTAACTTCAACCCAGACCAATTTGTGTGAGAGAGAGAGAGAGAG 1386  
DB 1420 GAAATCTTCTCTTAACTTCAACCCAGACCAATTTGTGTGAGAGAGAGAGAGAGAG 1479  
QY 1387 CCAAGAAAGCT 1446  
DB 1480 CCAAGAAAGCT 1539  
QY 1447 GTTCTGCGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1506  
DB 1540 GTTCTGCGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1599  
QY 1507 CCAAAATATCATGATCTCCATGAGAGACAGCGGAGAGAGAGAGAGAGAGAGAGAG 1566  
DB 1600 CCAAAATATCATGATCTCCATGAGAGACAGCGGAGAGAGAGAGAGAGAGAGAGAG 1659  
QY 1567 TCTCTCTGAAACAATTTCTTCTCACTTACAGATGAGAGAGAGAGAGAGAGAGAG 1626  
DB 1660 TCTCTCTGAAACAATTTCTTCTCACTTACAGATGAGAGAGAGAGAGAGAGAGAG 1719  
QY 1627 GTGGGTCCAGAAACCGGATAGCCCGTTTATT 1659  
DB 1720 GTGGGTCCAGAAACCGGATAGCCCGTTTATT 1752

RESULT 8  
US-10-741-600-693  
; Sequence 693, Application US/10741600  
; Publication No. US20050026169A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001499  
; CURRENT APPLICATION NUMBER: US/10/741,600  
; CURRENT FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 73997  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 693  
; LENGTH: 3274  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-741-600-693  
  
Query Match 42.0%; Score 879; DB 21; Length 3274;  
Best Local Similarity 99.1%; Pred. No. 0;  
Matches 1579; Conservative 0; Mismatches 14; Indels 0; Gaps 0;  
  
QY 67 TGTGAGCAAGCTGTGTGATCATGATTTTCTGCAATCTTCACTGATTAATGATTCGAT 126  
DB 178 TGTGAGCAAGCTGTGTGATCATGATTTTCTGCAATCTTCACTGATTAATGATTCGAT 237  
QY 127 AAGTATGACCTTAAACCGGAAACAGCTCTCTGTTGTGTGTGTTTCTTACACGAGGAC 186  
DB 238 AAGTATGACCTTAAACCGGAAACAGCTCTCTGTTGTGTGTGTTTCTTACACGAGGAC 297  
QY 187 GAGAGCCCAACCCGACAGAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACACTGCGG 246  
DB 298 GAGAGCCCAACCCGACAGAGCCCGCAAGTTTGTAAAGAAATACAGAACCAACACTGCGG 357  
QY 247 GTTGAATTTCTTCTCACTCTGCGGTATGAGTTTACGAGTCTGCGGTATTCAGATACCC 306  
DB 358 GTTGAATTTCTTCTCACTCTGCGGTATGAGTTTACGAGTCTGCGGTATTCAGATACCC 417  
QY 307 TACTTTTGAATGGGGGGAATTAATGATTAACGACTTCAAGAGCTTGAAGCCCGGACAT 366  
DB 418 TACTTTTGAATGGGGGGAATTAATGATTAACGACTTCAAGAGCTTGAAGCCCGGACAT 477  
QY 367 TTCTATGACACTGACATGACATGATGATGATGATGATGATGATGATGATGATGATGAT 426  
DB 478 TTCTATGACACTGACATGACATGATGATGATGATGATGATGATGATGATGATGATGAT 537  
QY 427 ATTGTGACATCTGACAGAGCTTCAAGAAAGCAATTTAAGTCAAGAGAGAGAGAGAGAG 486  
DB 538 ATTGTGACATCTGACAGAGCTTCAAGAAAGCAATTTAAGTCAAGAGAGAGAGAGAGAG 597  
QY 487 ATTAAGTGGGACATCTCCGCTGACATCACTGATCTCTTGAAGACAGACCTTGTGAAGTCA 546  
DB 598 ATTAAGTGGGACATCTCCGCTGACATCACTGATCTCTTGAAGACAGACCTTGTGAAGTCA 657  
QY 547 GAGCTGTACACATTAATCTCAAGTCAAGCTTCTGAGATTTGATGATGATGATGATGATG 606  
DB 658 GAGCTGTACACATTAATCTCAAGTCAAGCTTCTGAGATTTGATGATGATGATGATGATG 717  
QY 607 GATTCGAGTTTGAAGCAAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 666  
DB 718 GATTCGAGTTTGAAGCAAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 777  
QY 667 GACTTTGAGTCTCTCACTTACCCGTTGATGATGATGATGATGATGATGATGATGATGAT 726  
DB 778 GACTTTGAGTCTCTCACTTACCCGTTGATGATGATGATGATGATGATGATGATGATGAT 837  
QY 727 CCTGTTTACCCCGAATATTTCAGGATCATCTGACAGAGAGAGAGAGAGAGAGAGAGAG 786  
DB 838 CCTGTTTACCCCGAATATTTCAGGATCATCTGACAGAGAGAGAGAGAGAGAGAGAGAG 897  
QY 787 AGCCAAATATCTGTGATCTGACAGATCCAGTTTTCAGAGTCCAAATTTCAAGGCAAGTT 846  
DB 898 AGCCAAATATCTGTGATCTGACAGATCCAGTTTTCAGAGTCCAAATTTCAAGGCAAGTT 957  
QY 847 CAATCTACAGAAAG 906  
DB 958 CAATCTACAGAAAG 1017  
QY 907 ACAGACTTTTCTATCAGCTGAGAGATGCTTCAAGCTGATCTGCTTCAAGATGATTTCT 966



```
Db 1018 ACAGACTTTTCCATAGCTGAGAGATGCTTACAGGATGATGCTCCCTAAACATGATTTCT 1077
Qy 967 GAGGTACAAAGCTACTCCAAAGACTGACGCTTGAAGATTAAGAGAGAGACTGCTCTT 1026
Db 1078 GAGGTACAAAGCTACTCCAAAGACTGAGCTTGAAGATTAAGAGAGAGACTGCTCTT 1137
Qy 1027 TTGAAAAATTAAGGACAGACAAAGAAAGAAAGAGCTACCTTACCCTCAATATACCTGGG 1086
Db 1138 TTGAAAAATTAAGGACAGACAAAGAAAGAAAGAGCTACCTTACCCTCAATATACCTGGG 1197
Qy 1087 GAGTGTCTCTCAGTTCATTTTAACTGATGCTTGAATTCGAGAAATTCCTAAAAAG 1146
Db 1198 GAGTGTCTCTCAGTTCATTTTAACTGATGCTTGAATTCGAGAAATTCCTAAAAAG 1257
Qy 1147 GCATTTTTCGAGAGCCCTTGTGACATAACAGTACAGTCTGTAAGGACGAGCTACAG 1206
Db 1258 GCATTTTTCGAGAGCCCTTGTGACATAACAGTACAGTCTGTAAGGACGAGCTACAG 1317
Qy 1207 GAGCTGTGACGTAAACAAGGGGAGCCGATTAATACCGCTTTGTAAGATGCTGTGCC 1266
Db 1318 GAGCTGTGACGTAAACAAGGGGAGCCGATTAATACCGCTTTGTAAGATGCTGTGCC 1377
Qy 1267 TGCTTGTGATCTCCCTCGCTGCTTCCCTTCCCTTCCGACGACCACTCACTCTCTGCTC 1326
Db 1378 TGCTTGTGATCTCCCTCGCTGCTTCCCTTCCCTTCCGACGACCACTCACTCTCTGCTC 1437
Qy 1327 GAACATCTTCTTAACTTCAACCCAGACATATTCGTGACAGCTCAAGTATTATTTAC 1386
Db 1438 GAACATCTTCTTAACTTCAACCCAGACATATTCGTGACAGCTCAAGTATTATTTAC 1497
Qy 1387 CCAAGAAAGCTCAATTTGTCTTCAACATTTGTGAATTTGTCTTCTGACACACAGAG 1446
Db 1498 CCAAGAAAGCTCAATTTGTCTTCAACATTTGTGAATTTGTCTTCTGACACACAGAG 1557
Qy 1447 GTTCTGGGAGGAGATGATACAGGCTGCTGCTTGTGTTGTTCAATTTCTTCAAG 1506
Db 1558 GTTCTGGGAGGAGATGATGATACAGGCTGCTGCTTGTGTTGTTCAATTTCTTCAAG 1617
Qy 1507 CCAACATATGATGATCCATGAAAGACAGGAGGAGGAGCCCTGAGCTCTTAAGATATCCATC 1566
Db 1618 CCAACATATGATGATCCATGAAAGACAGGAGGAGGAGCCCTGAGCTCTTAAGATATCCATC 1677
Qy 1567 TCTCTCTGAAACAACAATTTCTTCACTTACAGATGACCCCTCAATCCCATCATATAG 1626
Db 1678 TCTCTCTGAAACAACAATTTCTTCACTTACAGATGACCCCTCAATCCCATCATATAG 1737
Qy 1627 GTGGGTCCAGGAACCGGATAGCCCGTTTATT 1659
Db 1738 GTGGGTCCAGGAACCGGATAGCCCGTTTATT 1770

RESULT 9
US-10-029-386-6369
; Sequence 6369, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 6369
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC008727.5
```

```
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.45
; OTHER INFORMATION: NT HIT: AF121205.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: Q9UPW0, EVALUE 3.30e+00
; OTHER INFORMATION: EST_HUMAN HIT: AU132586.1, EVALUE 0.00e+00
US-10-029-386-6369

Query Match 15.8%; Score 330; DB 16; Length 591;
Best Local Similarity 99.7%; Pred. No. 7,9e-170;
Matches 380; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 401 GTTTAGAACTTGTGTTGAGCCGTGATTTGCTGACCTCTGACAGCCCTCAGAAAGCATT 460
Db 38 GTTTAGAACTTGTGTTGAGCCGTGATTTGCTGACCTCTGACAGCCCTCAGAAAGCATT 97
Qy 461 TTAGGTCAAGCAGAGACAAAGAGAGATTAAGTGGCGCACTCCCGGTGGCATCCTGCAAT 520
Db 98 TTAGGTCAAGCAGAGACAAAGAGAGATTAAGTGGCGCACTCCCGGTGGCATCCTGCAAT 157
Qy 521 CCTGAGGACAGACCTTGTGAAGTCAAGCTGCTACATTTGAATTCGAAGTCAAGCTTTC 580
Db 158 CCTGAGGACAGACCTTGTGAAGTCAAGCTGCTACATTTGAATTCGAAGTCAAGCTTTC 217
Qy 581 TGAGATTGATGATTCAGGAGAAAGATTTGAGGTTTGAAGCAAAATGACGTACA 640
Db 218 TGAGATTGATGATTCAGGAGAAAGATTTGAGGTTTGAAGCAAAATGACGTACA 277
Qy 641 GCAACCAATCCAAATGTTGAATTAAGTCTGATGCTCACTTACCCGTTGGTACCC 700
Db 278 GCAACCAATCCAAATGTTGAATTAAGTCTGATGCTCACTTACCCGTTGGTACCC 337
Qy 701 CACTCTCAGAGCTTCTGATATTTCTGTTTACCCTCCAGAAATTTTACAGGTATC 760
Db 338 CACTCTCAGAGCTTCTGATATTTCTGTTTACCCTCCAGAAATTTTACAGGTATC 397
Qy 761 TGCAGAGTCTCTTGGCCAGG 781
Db 398 TGCAGAGTCTCTTGGCCAGG 418

RESULT 10
US-10-029-386-20100
; Sequence 20100, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 20100
; LENGTH: 379
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC008727.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.45
; OTHER INFORMATION: NT HIT: g114729757, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: Q9UPW0, EVALUE 1.80e+00
; OTHER INFORMATION: EST_HUMAN HIT: AU132586.1, EVALUE 0.00e+00
US-10-029-386-20100

Query Match 15.7%; Score 328; DB 16; Length 379;
Best Local Similarity 99.7%; Pred. No. 9,7e-169;
Matches 378; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 402 TTTAGAACTTGTGTTGAGCCGTGATTTGCTGACCTCTGACAGCCCTCAGAAAGCATT 461
Db 1 TTTAGAACTTGTGTTGAGCCGTGATTTGCTGACCTCTGACAGCCCTCAGAAAGCATT 60
```





QY	Db	QY	Db
702	301	702	301
ACCTCACAAGCCTCTCTGTAATTCCTGTTACCCCAAGATTTACAGGATCATCT	ACCTCACAAGCCTCTCTGTAATTCCTGTTACCCCAAGATTTACAGGATCATCT	ACCTCACAAGCCTCTCTGTAATTCCTGTTACCCCAAGATTTACAGGATCATCT	ACCTCACAAGCCTCTCTGTAATTCCTGTTACCCCAAGATTTACAGGATCATCT
762	361	762	361
GCAGGATCTCTTTGGCCAG	GCAGGATCTCTTTGGCCAG	GCAGGATCTCTTTGGCCAG	GCAGGATCTCTTTGGCCAG
761	360	761	360
ACCTCACAAGCCTCTCTGTAATTCCTGTTACCCCAAGATTTACAGGATCATCT	ACCTCACAAGCCTCTCTGTAATTCCTGTTACCCCAAGATTTACAGGATCATCT	ACCTCACAAGCCTCTCTGTAATTCCTGTTACCCCAAGATTTACAGGATCATCT	ACCTCACAAGCCTCTCTGTAATTCCTGTTACCCCAAGATTTACAGGATCATCT

RESULT 13  
US-10-741-600-17757

```

: Sequence 17757, Application US/10741600
: Publication No. US20050026169A1
:
: GENERAL INFORMATION:
: APPLICANT: CARGILL, Michele et al.
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: CLO01499
: CURRENT APPLICATION NUMBER: US/10/741,600
: CURRENT FILING DATE: 2003-12-22
: NUMBER OF SEQ ID NOS: 73997
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 17757
:   LENGTH: 43985
:   TYPE: DNA
: ORGANISM: Homo sapiens
: US-10-741-600-17757

```

Query Match	12.7%	Score 266;	DB 21;	Length 43985;
Best Local Similarity	99.5%;	Pred. No. 1.4e-134;		
Matches 366;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;

Qy	401	GTTTAGAACTTGTGGTTGAGCCCGGAGTTGCTGAACTCTGGACAGCCCTCAGAAAGCATT	460
Db	14836	GTTTAGAACTTGTGGTTGAGCCCGGAGTTGCTGAACTCTGGACAGCCCTCAGAAAGCATT	14895
Qy	461	TTAGGTCAGACAGGACAAAGAGAGATTAAGTGGCCGACTCCCGGTGGCAATCACTTGCAT	520
Db	14896	TTAGGTCAGACAGGACAAAGAGAGATTAAGTGGCCGACTCCCGGTGGCAATCACTTGCAT	14955
Qy	521	CCTTGAGAGACAGACTTGTGAAGTCAGAGCTGGTACACATTTGAATCTCAAGTCGAGCTTC	580
Db	14956	CCTTGAGAGACAGACTTGTGAAGTCAGAGCTGGTACACATTTGAATCTCAAGTCGAGCTTC	15015
Qy	581	TGAGATTCGATGATTCAGAAAGAAAGATTCTGAGGTTTGAAGCAAAATGCAAGTGACACA	640
Db	15016	TGAGATTCGATGATTCAGAAAGAAAGATTCTGAGGTTTGAAGCAAAATGCAAGTGACACA	15075
Qy	641	GCAACCAATCAATGTTGTAAATTGGAAGACTTGAAGTCCACATTACCCGTTGGTACCCCC	700
Db	15076	GCAACCAATCAATGTTGTAAATTGGAAGACTTGAAGTCCACATTACCCGTTGGTACCCCC	15135
Qy	701	CACCTCTCAAGAGCCCTCTGATATATCTCTGGTTTACCCCCAGAAATTTTACAGGTACATC	760
Db	15136	CACCTCTCAAGAGCCCTCTGATATATCTCTGGTTTACCCCCAGAAATTTTACAGGTACATC	15195
Qy	761	TGCAGAGAG 768	
Db	15196	TGCAGAGAG 15203	

RESULT 14  
US-10-029

Sequence 633 Application US/10029386  
Publication No. US20030194704A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Hank, David R.  
APPLICANT: Hank, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
TITLE OF INVENTION: EXPRESSION ANALYSIS TWO  
FILE REFERENCE: AEOmica-X-2

```

; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1.1
; SEQ ID NO: 333

```

```

? SEQ ID NO 633
? LENGTH: 525
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? OTHER INFORMATION: MAP TO AC021609.3
? OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.48
? OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.58
? OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.52
? OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.57
? OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.79
? OTHER INFORMATION: SWISSPROT HIT: P37039, EVALUATE 1.00e-06
? OTHER INFORMATION: EST HUMAN HIT: BF346446.1, EVALUATE 1.00e-58
? OTHER INFORMATION: NT HIT: AF121212.1, EVALUATE 0.00e+00
US-10-029-386-633

```

Query Match	9.0%	Score 188;	DB 16;	length 525;
Best Local Similarity	100.0%	Pred. No. 8.7e-92;		
Matches 188; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1761	TTGGAAGAAAGGCTCAGACATTTCTCTTAAGCATGGATCTTAATCATCTAAAGTTTCC	1820
QY	234	TTGAGAAAGAGCTCAGACATTTCTTAAGCATGGATCTTAATCATCTAAAGTTTCC	175
QY	1821	TTCTCAAGAGATGCTCCTGTGGGAGAGAGAACCCACGAAAGTATGTACAGACAAAC	1880
Db	174	TTCTCAAGAGATGCTCCTGTGGGAGAGAGAACCCACGAAAGTATGTACAGACAAAC	115
QY	1881	ATTCAGCTTCAATGGCCAGCAGGTGGCGAAGATCTCTCTCCAGAGAAACGGCCATATTTAT	1940
Db	114	ATTCAGCTTCAATGGCCAGCAGGTGGCGAAGATCTCTCTCCAGAGAAACGGCCATATTTAT	55
QY	1941	GTGTGTGG 1948	
Db	54	GTGTGTGG 47	

RESULT 15  
US-10-029-386-14338/c  
; Sequence 14338, Application US/10029386  
; Publication No. US20030194704A1

```

: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GR
: TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
: FILE REFERENCE: ABOICA-X-2
: CURRENT APPLICATION NUMBER: US/10/029,386
: CURRENT FILING DATE: 2001-12-20
: NUMBER OF SEQ ID NOS: 34288
: SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 14338
: LENGTH: 175
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AC021609.3
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.48
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.58
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.52
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.57
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.79
: OTHER INFORMATION: SWISSPROT HIT: O61608, EVALUE 4.00e-04
: OTHER INFORMATION: EST HUMAN HIT: AA085543.1, EVALUE 7.00e-94
: OTHER INFORMATION: NT HIT: G113325067, EVALUE 5.00e-94
: US-10-029-386-14338

```

Query Match	8.48;	Score 175;	DB 16;	Length 175;
-------------	-------	------------	--------	-------------

Best Local Similarity 100.0%; Pred. No. 1,1e-84;  
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1766 AAAAGAGCTCAGACATTTCTTAAGCATGGATCTTAACCTCATCTAAAGGTTTCCTTCTC 1825

Db 175 AAAAGAGCTCAGACATTTCTTAAGCATGGATCTTAACATCATCTAAAGGTTCCCTTCTC 116

1826 AAGAGATGCTCTGTTGGGAGGAGGAAGCCCAAGCAAGTATGTACAAGACAACATCCA 1885

Db 115 AAGAGATGCTCCTGTTGGGAGGAGGAGCCCGAGCAAGTATGTACAGAACAATCCA 56

1886 GCTTCATGGCCAGCAGGTGGCGAATCTCTCCAGGAAACGCCATATTAT 1940

Db 55 GCTTCATGGCCAGCAGGTGGCGGAATCCTCCTCCAGGAGAACGGCCATATTAT 1

Search completed: August 27, 2005, 17:33:35  
Job time : 901.684 secs

Job time : 901.684 secs

**This Page Blank (uspto)**